International Adult Literacy Survey

Microdata User's Guide

Statistics Canada Statistique Canada

Canada

Table of Contents

The blue highlighted texts are links to the corresponding pages. Activate the hand tool, point to the blue highlighted text and click.

To view any of the questionnaires or record layouts, activate the hand tool and click on the corresponding country name. Once the questionnaire or record layout is opened, you can link back to the user's guide by clicking inside the blue box on the first or last page of the selected document.

1.0 Introduction

The Participants

The International Adult Literacy Survey (IALS) was a large-scale co-operative effort by governments, national statistical agencies, research institutions and the Organisation for Economic Co-operation and Development (OECD). The development and management of the survey were co-ordinated by Statistics Canada and the Educational Testing Service of Princeton, New Jersey. At various survey cycles, and in different ways, substantial input was received from the National Center for Education Statistics of the United States Department of Education, input that has greatly facilitated the project and ultimately made the following publications possible.

In 1994, nine countries – Canada (English and French-speaking populations), France, Germany, Ireland, the Netherlands, Poland, Sweden, Switzerland (German and French-speaking regions) and the United States – fielded the world's first large-scale, comparative assessment of adult literacy. Data for seven of these countries were published in *Literacy, Economy and Society: Results of the First International Adult Literacy Survey* in December 1995 (OECD and Statistics Canada, 1995).[1]

Encouraged by this demonstration of success, five additional countries or territories – Australia, the Flemish Community in Belgium, Great Britain, New Zealand and Northern Ireland – decided to administer the IALS instruments to samples of their adult populations in 1996. Comparative data from this round of collection were released in November 1997 in *Literacy Skills for the Knowledge Society: Further Results from the International Adult Literacy Survey* (OECD and HRDC, 1997).

Nine other countries or regions – Chile, the Czech Republic, Denmark, Finland, Hungary, Italy, Norway, Slovenia and the Italian-speaking region of Switzerland – participated in a third, large-scale round of data collection in 1998. Results for most of these countries are included in the publication entitled *Literacy in the Information Age: Final Report of the International Adult Literacy Survey* (OECD and Statistics Canada, 2000). Japan, Malaysia, Mexico and the Canary Islands region of Spain have also successfully experimented with IALS-derived instruments.[2]

Participants of the first two rounds of IALS are referenced in this guide as part of the first cycle, whereas participants of the third round are referenced as the second cycle.

1 France decided to withdraw from the study in November 1995, citing concerns over comparability. Data processing for Ireland was unfortunately delayed and so its results were included in a subsequent IALS publication.

2 Results for these countries are not included in this report because they were obtained in feasibility studies that used limited and non-representative samples.

The data file in this package is a compilation of the IALS datasets received from the various participating countries. No changes to these datasets have been made from what was received from each country. The documentation for individual countries that is provided in this manual is the information that was provided by each IALS country. Further information on the individual data files or supporting documentation should thus be addressed to the appropriate study manager. Their contact information is given below. It should be noted that Australian IALS data is only available through the Australian Bureau of Statistics, for confidentiality reasons.

Several countries have published National Reports as well — the respective National study managers as outlined below should be contacted for additional details.

This document summarizes the survey concepts and operations of the international survey. It is important for users to become familiar with the contents of this document before publishing or otherwise releasing any estimates derived from the IALS microdata file.

National Study Managers

Australia
Mr. Mel Butler
Australian Bureau of Statistics
National Centre for Education and Training Statistics
P.O. Box 10
Belconnen
ACT 2616
Australia
Tel: +61 2 6252 5936
Fax: +61 2 6252 8013
Internet: mel.butler@abs.gov.au

Belgium
Mr. Luc van de Poele
Universiteit Gent
Henri Dunantlaan 2, 9000 Gent
Belgium
Tel: +32 9 264 6398
Fax: +32 9 233 1098
Internet: luc.vandepoele@rug.ac.be

Canada
Mr. Jean Pignal
Special Surveys Division, Statistics Canada
5-A5, Jean Talon Bldg.
Tunney's Pasture, Ottawa, Ontario K1A 0T6
Tel: 613-951-3317
Fax: 613-951-0562
Internet: pignjea@statcan.ca

Chile

Mr. David Bravo
Departamento de Economia
Facultad de Ciencias Economicas y Admin.
Universidad de Chile
Diagonal Paraguay #257, Torre 26, Oficina 1503
Santiago, Chile
Tel: +56 2 678 3412
Fax: +56 2 678 3413
Internet: dbravo@decon.facea.uchile.cl

Czech Republic

Dr. Petr Mateju
Institute of Sociology Academy of Sciences of the Czech Republic
Jilska 1
110 00 Prague 1
Czech Republic
Tel: +420 2 2222 0098 - 101
Fax: +420 2 2222 0143
Internet: mateju@mbox.cesnet.cz

Denmark

Mr. Torben Pilegaard Jensen
AKF
Nyropsgade 37
1602 Copenhagen V
Denmark
Tel: +45 33 11 0300
Fax: +45 33 15 2875
Internet: tpj@akf.dk

Finland

Ms. Pirjo Linnakylä
Institute for Educational Research
University of Jyväskylä
P.O. Box 35
FIN-40351 Jyväskylä
Finland
Tel: +358 14 603 281
Fax: +358 14 603 201
Internet: linnakyl@jvu.fi

Germany

Mr. Rainer Lehmann
Humboldt-Universitat Zu Berlin
Institut fur Allgemeine Padagogik
Unter Den Linden 6
10099 Berlin, Germany
Tel: +49 30 2093 4132
Fax: +49 30 2093 4153
Internet: rainer.lehmann@educat.hu-berlin.de

Hungary

Ms. Judit Krolopp
National Institute of Public Education
Centre for Evaluation Studies
PO Box 120
1364 Budapest
Hungary
Tel: +36 1 318 0865
Fax: +36 1 318 6143
Internet: krolopp@ppp.ces.hu

Ireland

Mr. Mark Morgan
Educational Research Centre
St. Patrick's College, Dublin 9, Ireland
Tel: +353 1 37 37 89
Fax: +353 1 837 8987
Internet: David@erc.ie

Italy

Ms. Vittoria Gallina
Centro Europeao Dell'Educazione
CEDE Villa Falconieri, Frascati
Roma 00044
Italy
Tel: +39 06 941 851
Fax: +39 06 941 85 205
Internet: vgallina@cede.it

Netherlands

Mr. Willem Houtkoop
Universiteit van Amsterdam
Faculteit Maatschappij en Gedragswetsenschappen
Max Goote Kenniscentrum
Wibautstraat 4
1091 GM Amsterdam
Netherlands
Tel: +31 020 525 1245
Fax: +31 020 525 1270
Internet: willem@educ.uva.nl

New Zealand

Ms. Lynne Whitney
Senior Manager, Research Division
Ministry of Education
PO Box 1666
Wellington, New Zealand
Tel: +64 4 463 8310
Internet : lynne.whitney@minedu.govt.nz

Norway

Mr. Egil Gabrielsen
Centre for Reading Research
Box 2504, Ullandhaug
4091 Stavanger
Norway
Tel: +4751 83 3200
Fax: +4751 83 3250
Internet: egil.gabrielsen@slf.his.no

Poland

Mr. Ireneusz Bialecki
Warsaw University
00-46 Warsaw, ul Nowt Swiat 69
Poland
Tel: +48 22 26 07 461
Fax: Same as telephone number
Internet: crphe@plearn.binet

Slovenia

Ms. Ester Mozina
Slovene Adult Education Centre
Smartinska 134a
1000 Ljubljana
Slovenia
Tel: +386 61 1842 591
Fax: +386 61 445 881
Internet: ester.mozina@acs-saec.si

Sweden

Mr. Mats Myberg
Linköping University
S-58183 Linköping
Sweden
Tel: +46 13 282 109
Fax: +46 13 282 145
Internet: matmy@IPP.LIU.se

Switzerland

Mr. Francois Stoll (French survey)
Institute for Applied Psychology
University of Zurich, Schonbergasse 2, Zurich
Switzerland
Tel: +41 1 634 3741
Fax: +41 1 634 6953
Internet: fstoll@angpsy.unizh.ch

Mr. Philipp Notter (German survey)
Kompetenzzentrum fur Bildungsevaluation und Leistungsmessung
Seilergraben 53
CH-8001 Zurich
Switzerland
Tel: +41 1 634 3580
Fax: +41 1 634 3587
Internet: Philipp.Notter@access.unizh.ch

Ms. Francesca Pedrazzini-Pesce (Italian survey)
Ufficio Studi e Ricerche
Stabile Torretta
6501 Bellinzona
Switzerland
Tel: +41 91 814 3491/18
Fax: +41 91 814 4436
Internet: francesca.pedrazzini-pesce@ti.ch

United Kingdom

Great Britain Ms. Siobhan Carey
National Office of Population Statistics
1 Drummond Gate, Pimlico, Great Britain
London SW1V2QQ
Tel: +44 171 533 5371
Fax: +44 171 533 5300
Internet: Siobhan.Carey@ons.gov.uk

Northern Ireland Mr. Kevin Sweeney
Central Survey Unit
Londonderry House, Chichester Street
Belfast BT14SX, Northern Ireland
Tel: +44 1232 252 555
Fax: +44 1232 252 534
Internet: kevin.sweeney.csu@nics.gov.uk

United States Dr. Marilyn Binkley
National Center for Education Statistics
1990 K Street North West
Washington, D.C. 20006
Tel: 202-219-2195
Fax: 202-501-3005
Internet: marilyn_binkley@ed.gov

For more information on this microdata package, please contact Statistics Canada's Centre for Education Statistics at: educationstats@statcan.ca or call 1-800-307-3382 (North America only) or (613) 951-7608 or Fax: (613) 951-9040.

2.0 Background

In recent years, adult literacy has come to be seen as crucial to the economic performance of industrialized nations. Literacy is no longer defined merely in terms of a basic threshold of reading ability which everyone growing up in developed countries is expected to attain. Rather, literacy is now equated with an individual's ability to use written information to function in society. Unlike their predecessors, adults today need a higher level of literacy to function well, because society has become more complex and low-skill jobs are disappearing. Inadequate levels of literacy in a broad section of the population may therefore have serious implications, even threatening a nation's economic strength and social cohesion.

Because of these high stakes, governments have a growing interest in understanding the level and distribution of literacy within their adult populations, and in learning what can be done to improve literacy. Accordingly, in recent years, many governments have tried for the first time to measure adult literacy directly. Pioneering studies (Kirsch and Jungeblut 1986; Kirsch and Mosenthal 1990; Statistics Canada 1991; Kirsch, Jungeblut, and Campbell 1992; Kirsch, Jungeblut, Jenkins, and Kolstad 1993) published in North America in the early 1990s revealed that significant percentages of adults lacked the literacy skills they were likely to need in everyday life. In 1992, the Organisation for Economic Co-operation and Development (OECD) concluded that low literacy levels were a serious threat to economic performance and social cohesion (OECD 1992). Yet a lack of comparable international data prevented a broader inquiry into literacy problems and consequent policy lessons across industrialized countries.

The International Adult Literacy Survey (IALS) was undertaken by 20 governments[3] and three intergovernmental organizations[4] in a collaborative effort to fill this need for information. In this survey, large samples of adults (ranging from 1,500 to 6,000 per country) worldwide were given the same broad test of their literacy skills during between 1994 and 1998. The results provide the most detailed portrait ever created on the condition of adult literacy and its relationship with an array of background and demographic characteristics.

3 Australia, Belgium, Canada, Chile, Czech Republic, Denmark, Finland, Germany, Hungary, Ireland, Italy, Netherlands, New Zealand, Norway, Poland, Slovenia, Sweden, Switzerland (German, French, Italian), United Kingdom (Great Britain, Northern Ireland) and United States.
4 OECD, European Union, and UNESCO.

3.0 Objectives

The IALS venture was initiated with two fundamental goals:

1) The first objective was to develop an assessment instrument that would permit useful comparisons of literacy performance across languages and cultures.

2) If such an assessment could be created, the second goal was to perform such comparisons, describing the literacy skills of people from different countries; each country's skill profile would be obtained by conducting a sample survey of households representative of the entire adult population.

The central element of the survey was the direct assessment of the literacy skills of respondents using commonplace tasks of varying degree of difficulty drawn from a range of topic and knowledge areas. This information was supported by the collection of background information on respondents. In addition, the background questionnaire included questions on the self-assessment of literacy skills of respondents, on the training which the respondent has taken in the year previous to the survey and on the perceived barriers to realizing enhanced literacy skill levels.

4.0
Concepts and Definitions

This chapter outlines concepts and definitions of interest to the users. Users are referred to section 9 of this document for a copy of the actual survey forms used.

4.1
Defining and Measuring Literacy

Many studies have treated literacy as a condition that adults either have or do not have, and thereby tried to count the number of illiterate members of the population. Such efforts typically define literacy in terms of the number of years of schooling completed, or by grade-level scores on school-based reading tests.

The IALS survey design team agreed that it would be undesirable to establish a single international standard for literacy. Such a standard would not only be arbitrary, but would also fail to acknowledge the multifaceted nature of literacy and the complexity of the literacy problem. Therefore, the participating countries concurred that, in common with recent North American and Australian surveys, the IALS would define literacy in terms of a mode of adult behaviour, namely:

Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.

This definition attempts to encompass a broad set of information-processing skills that adults may use in performing different types of tasks at work, at home, or in their communities. Some other types of knowledge and skill (including teamwork, interpersonal skills, and other communication skills) were also recognized as being important but could not be measured with the resources available.

According to the IALS definition, literacy is neither a single skill used in dealing with all types of text, nor an infinite set of skills, each particular to a different type of material. Thus, following the example of the North American studies noted earlier, the IALS team defined three domains of literacy:

a) *Prose literacy*—the knowledge and skills needed to understand and use information from texts including editorials, news stories, poems, and fiction;
b) *Document literacy*—the knowledge and skills required to locate and use information contained in various formats, including job applications, payroll forms, transportation schedules, maps, tables, and graphics; and
c) *Quantitative literacy*—the knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials, such as balancing a checkbook, calculating a tip, completing an order form, or determining the amount of interest on a loan from an advertisement.

Rather than define a threshold for competency—a standard that distinguishes the so-called "literate" from the "illiterate"—researchers constructed a scale from 0 to 500 in each of these three literacy domains. Tasks of varying difficulty can be placed along these scales. A person's literacy ability in each domain is expressed by a score, which is the point on the scale at which he or she has an 80 per cent chance of successfully performing a given literacy task. Individuals can then be grouped into five levels of literacy, defined by score ranges (Level 1 includes scores from 0 to 225, Level 2 contains scores from 226 to 275, and so on). These levels are useful in analyzing and reporting the survey results and in designing remedial programs.

The use of three parallel literacy scales makes it possible to profile and compare the various types and levels of literacy demonstrated by adults in different countries and by subgroups within those countries. The scales also help policy makers, business leaders, educators, and others to understand the broad and diverse nature of literacy.

4.2

Defining and Measuring Literacy Performance on Three Scales

The results of the International Adult Literacy Survey (IALS) are reported on three scales – prose, document and quantitative – rather than on a single scale. Each scale ranges from 0 to 500. Scale scores have, in turn, been grouped into five empirically determined literacy levels. Each of these levels implies an ability to cope with a particular subset of reading tasks. This section explains in more detail how the proficiency scores can be interpreted, by describing the scales and the kinds of tasks that were used in the test, and the literacy levels that have been adopted.[5]

While the literacy scales make it possible to compare the prose, document and quantitative skills of different populations and to study the relationships between literacy skills and various factors, the scale scores by themselves carry little or no meaning. In other words, whereas most people have a practical understanding of what it means when the temperature outside reaches 10°C, it is not intuitively clear what it means when a particular group is at 287 on the prose scale, or 250 on the document scale, or at level 2 on the quantitative scale.

One way to gain some understanding about what it means to perform at a given point along a literacy scale is to identify a set of variables that can be shown to underlie performance on these tasks. Collectively, these variables provide a framework for understanding what is being measured in a particular assessment, and what knowledge and skills are being demonstrated by various levels of proficiency.

Toward this end, the following text begins by describing how the literacy scale scores were defined. Detailed descriptions of the prose, document and quantitative scales are then provided, along with definitions of the five levels. Sample tasks are presented to illustrate the types of materials and task demands that characterise the levels.

5 This text is partially reprinted from Chapter 2 in Literacy, Economy and Society (OECD and Statistics Canada, 1995).

4.3

Defining the Literacy Levels

The Item Response Theory (IRT) scaling procedures that were used in the IALS constitute a statistical solution to the challenge of establishing one or more scales for a set of tasks with an ordering of difficulty that is essentially the same for everyone.[6] First, the difficulty of tasks is ranked on the scale according to how well respondents actually perform them. Next, individuals are assigned scores according to how well they perform on a number of tasks of varying difficulty.

The scale point assigned to each task is the point at which individuals with that proficiency score have a given probability of responding correctly. In this survey, an 80 per cent probability of correct response was the criterion used. This means that individuals estimated to have a particular scale score will perform tasks at that point on the scale with an 80 per cent probability of a correct response. It also means they will have a greater than 80 per cent chance of performing tasks that are lower on the scale. It does not mean, however, that individuals with given proficiencies can never succeed at tasks with higher difficulty values; they may do so some of the time. It does suggest that their probability of success is "relatively" low – i.e. the more difficult the task relative to their proficiency, the lower the likelihood of a correct response.

An analogy might help clarify this point. The relationship between task difficulty and individual proficiency is much like the high jump event in track and field, in which an athlete tries to jump over a bar that is placed at increasing heights. Each high jumper has a height at which he or she is proficient – that is, the jumper can clear the bar at that height with a high probability of success, and can clear the bar at lower heights almost every time. When the bar is higher than the athlete's level of proficiency, however, it is expected that the athlete will be unable to clear the bar consistently.

Once the literacy tasks are placed along each of the scales using the criterion of 80 per cent, it is possible to see to what extent the interactions among various task characteristics capture the placement of tasks along the scales. Analyses of the task characteristics which include the materials being read and the type of questions asked about these materials reveal that ordered sets of information-processing skills appear to be called into play to successfully perform the various tasks displayed along each scale (Kirsch and Mosenthal, 1993).

To capture this order, each scale is divided into five levels reflecting the empirically determined progression of information-processing skills and strategies. While some of the tasks were at the low end of a scale and some at the very high end, most had values in the 200-to-400 range. It is important to recognise that these levels were selected not as a result of any inherent statistical property of the scales, but rather as the result of shifts in the skills and strategies required to succeed at various tasks along the scales, ranging from simple to complex.

6 The reader is referred to Murray, Kirsch and Jenkins (1997) for a complete description of the scaling procedures used in this assessment.

The remainder of this section describes each scale in terms of the nature of task demands at each of the five levels. Sample tasks are presented and the factors contributing to their difficulty discussed. The aim is to facilitate interpretation of the results and data analyses.

4.4

Interpreting the Literacy Levels

This section describes each scale in terms of the nature of task demands at each of the five levels. For each scale, the factors contributing to their difficulty are discussed. The aim of the section is to provide meaning to the scales and to facilitate interpretation of the results.

4.4.1

Prose Literacy

The ability to understand and use information contained in various kinds of text is an important aspect of literacy. The study therefore included an array of prose selections, including text from newspapers, magazines and brochures. The material varied in length, density of text, content, and the use of structural or organisational aids such as headings, bullets and special typefaces. All prose samples were reprinted in their entirety with the original layout and typography unchanged.

Each prose selection was accompanied by one or more questions asking the reader to perform specific tasks. These tasks represent three major aspects of information-processing: locating, integrating and generating. Locating tasks require the reader to find information in the text based on conditions or features specified in the question or directive. The match may be literal or synonymous, or the reader may need to make an inference in order to perform successfully. Integrating tasks ask the reader to pull together two or more pieces of information in the text. The information could be found in a single paragraph, or in different paragraphs or sections. With the generating tasks, readers must produce a written response by processing information from the text and by making text-based inferences or drawing on their own background knowledge.

In all, the prose literacy scale includes 34 tasks with difficulty values ranging from 188 to 377. These tasks are distributed by level as follows: Level 1, 5 tasks; Level 2, 9 tasks; Level 3, 14 tasks; Level 4, 5 tasks; and Level 5, 1 task. It is important to remember that the tasks requiring the reader to locate, integrate and generate information extend over a range of difficulty as a result of combining other variables, including:

- the number of categories or features of information the reader must process;
- the extent to which information given in the question or directive is obviously related to the information contained in the text;
- the amount and location of information in the text that shares some of the features with the information being requested and thus appears relevant, but that in fact does not fully answer the question (these are called “distractors”);
- the length and density of the text.

The five levels of prose literacy are defined as follows.

Prose level 1 Score range: 0 to 225

Most of the tasks at this level require the reader to locate one piece of information in the text that is identical to or synonymous with the information given in the directive. If a plausible incorrect answer is present in the text, it tends not to be near the correct information.

Typically the match between the task and the text is literal, although sometimes a low-level inference may be necessary. The text is usually brief or has organisational aids such as paragraph headings or italics that suggest where the reader can find the specified information. Generally, the target word or phrase appears only once in the text.

The easiest task in level 1 (difficulty value of 188) directs respondents to look at a medicine label to determine the "maximum number of days you should take this medicine". The label contains only one reference to number of days and this information is located under the heading "DOSAGE". The reader must go to this part of the label and locate the phrase "not longer than 7 days".

MEDCO ASPIRIN *500*

INDICATIONS: Headaches, muscle pains, rheumatic pains, toothaches, earaches. RELIEVES COMMON COLD SYMPTOMS.

DOSAGE: ORAL. 1 or 2 tablets every 6 hours, preferably accompanied by food, for not longer than 7 days. Store in a cool, dry place.

CAUTION: Do not use for gastritis or peptic ulcer. Do not use if taking anticoagulant drugs. Do not use for serious liver illness or bronchial asthma. If taken in large doses and for an extended period, may cause harm to kidneys. Before using this medication for chicken pox or influenza in children, consult with a doctor about Reyes Syndrome, a rare but serious illness. During lactation and pregnancy, consult with a doctor before using this product, especially in the last trimester of pregnancy. If symptoms persist, or in case of an accidental overdose, consult a doctor. Keep out of reach of children.

INGREDIENTS: Each tablet contains
500 mg acetylsalicicylic acid.
Excipient c.b.p. 1 tablet.
Reg. No. 88246

Made in Canada by STERLING PRODUCTS, INC.
1600 Industrial Blvd., Montreal, Quebec H9J 3P1

0 67736 11079

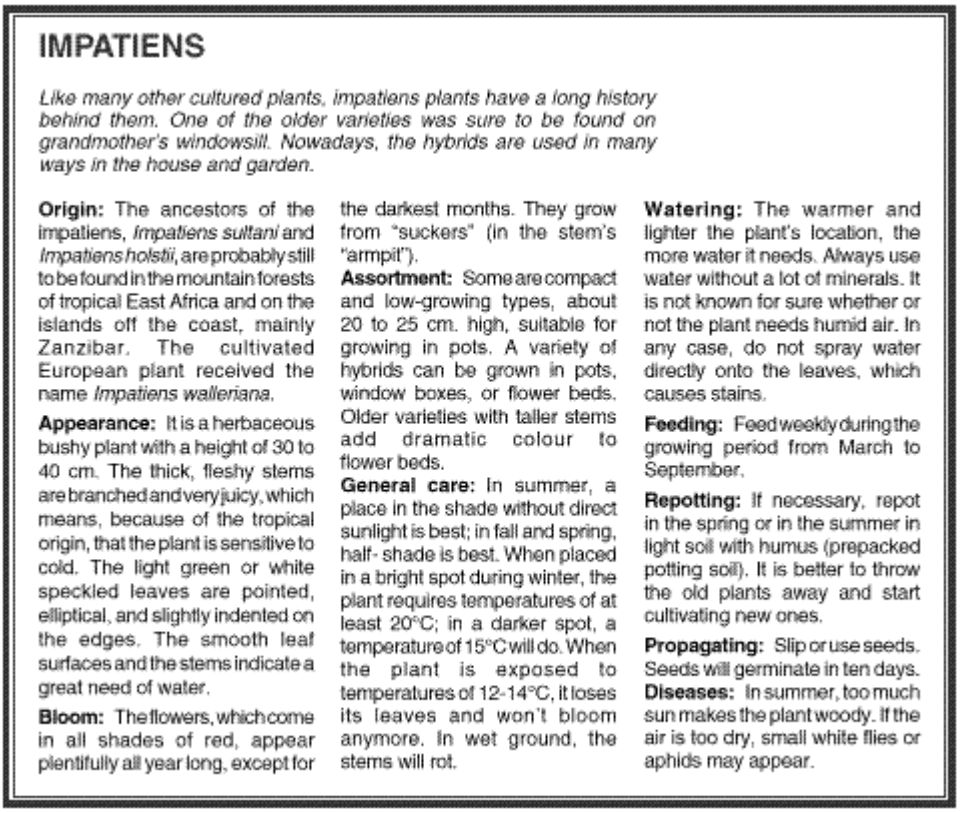
Reprinted by permission

Prose level 2 Score range: 226 to 275

Tasks at this level generally require the reader to locate one or more pieces of information in the text, but several distractors may be present, or low-level inferences may be required. Tasks at this level also begin to ask readers to integrate two or more pieces of information, or to compare and contrast information.

As with level 1, most of the tasks at level 2 ask the reader to locate information. However, more varied demands are made in terms of the number of responses the question requires, or in terms of the distracting information that may be present. For example, a task based on an article about the impatiens plant asks the reader to determine what happens when the plant is exposed to temperatures of 14°C or lower. A sentence under the section "General care" states that "When the plant is exposed to temperatures of 12-14°C, it loses its leaves and won't bloom anymore." This task received a difficulty value of 230, just in the level 2 range.

What made this task somewhat more difficult than those identified at level 1 is that the previous sentence in the text contains information about the requirements of the impatiens plant in various temperatures. This information could have distracted some readers, making the task slightly more difficult.



IMPATIENS

Like many other cultured plants, impatiens plants have a long history behind them. One of the older varieties was sure to be found on grandmother's windowsill. Nowadays, the hybrids are used in many ways in the house and garden.

Origin: The ancestors of the impatiens, *Impatiens sultani* and *Impatiens holstii*, are probably still to be found in the mountain forests of tropical East Africa and on the islands off the coast, mainly Zanzibar. The cultivated European plant received the name *Impatiens walleriana*.

Appearance: It is a herbaceous bushy plant with a height of 30 to 40 cm. The thick, fleshy stems are branched and very juicy, which means, because of the tropical origin, that the plant is sensitive to cold. The light green or white speckled leaves are pointed, elliptical, and slightly indented on the edges. The smooth leaf surfaces and the stems indicate a great need of water.

Bloom: The flowers, which come in all shades of red, appear plentifully all year long, except for the darkest months. They grow from "suckers" (in the stem's "armpit").

Assortment: Some are compact and low-growing types, about 20 to 25 cm. high, suitable for growing in pots. A variety of hybrids can be grown in pots, window boxes, or flower beds. Older varieties with taller stems add dramatic colour to flower beds.

General care: In summer, a place in the shade without direct sunlight is best; in fall and spring, half- shade is best. When placed in a bright spot during winter, the plant requires temperatures of at least 20°C; in a darker spot, a temperature of 15°C will do. When the plant is exposed to temperatures of 12-14°C, it loses its leaves and won't bloom anymore. In wet ground, the stems will rot.

Watering: The warmer and lighter the plant's location, the more water it needs. Always use water without a lot of minerals. It is not known for sure whether or not the plant needs humid air. In any case, do not spray water directly onto the leaves, which causes stains.

Feeding: Feed weekly during the growing period from March to September.

Repotting: If necessary, repot in the spring or in the summer in light soil with humus (prepacked potting soil). It is better to throw the old plants away and start cultivating new ones.

Propagating: Slip or use seeds. Seeds will germinate in ten days.

Diseases: In summer, too much sun makes the plant woody. If the air is too dry, small white flies or aphids may appear.

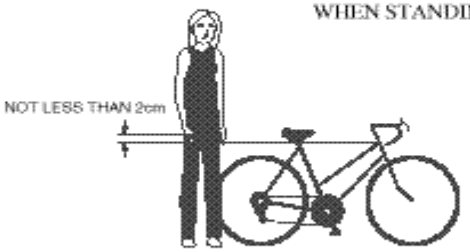
A similar task involving the same text asks the reader to identify "what the smooth leaf and stem suggest about the plant". The second paragraph of the article is labelled "Appearance" and contains a sentence that states, ". . . stems are branched and very juicy, which means, because of the tropical origin, that the plant is sensitive to cold." This sentence distracted some readers from the last sentence in that same paragraph: "The smooth leaf surfaces and the stems indicate a great need of water." This task received a difficulty value of 254, placing it in the middle of level 2.

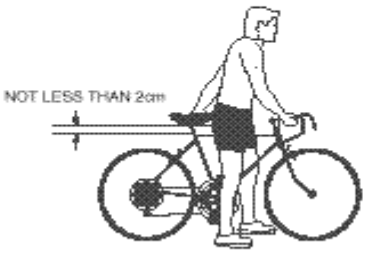
Prose level 3 Score range: 276 to 325

Tasks at this level generally direct readers to locate information that requires low-level inferences or that meets specified conditions. Sometimes the reader is required to identify several pieces of information that are located in different sentences or paragraphs rather than in a single sentence. Readers may also be asked to integrate or to compare and contrast information across paragraphs or sections of text.

PROPER FRAME FIT

RIDER MUST BE ABLE TO STRADDLE BICYCLE WITH AT LEAST 2 cm CLEARANCE ABOVE THE HORIZONTAL BAR WHEN STANDING.





NOTE: Measurement for a female should be determined using a men's model as a basis.

PROPER SIZE OF BICYCLE	
FRAME SIZE	LEG LENGTH OF RIDER
430mm	660mm-760mm
460mm	690mm-790mm
480mm	710mm-790mm
530mm	760mm-840mm
560mm	790mm-860mm
580mm	810mm-890mm
635mm	860mm-940mm

OWNER'S RESPONSIBILITY

1. **Bicycle Selection and Purchase:** Make sure this bicycle fits the intended rider. Bicycles come in a variety of sizes. Personal adjustment of seat and handlebars is necessary to assure maximum safety and comfort. Bicycles come with a wide variety of equipment and accessories . . . make sure the rider can operate them.
2. **Assembly:** Carefully follow all assembly instructions. Make sure that all nuts, bolts and screws are securely tightened.
3. **Fitting the Bicycle:** To ride safely and comfortably, the bicycle must fit the rider. Check the seat position, adjusting it up or down so that with the sole of rider's foot on the pedal in its lowest position the rider's knee is slightly bent.

Note: Specific charts illustrated at left detail the proper method of determining the correct frame size.

The manufacturer is not responsible for failure, injury, or damage caused by improper completion of assembly or improper maintenance after shipment.

One level 3 task (with a difficulty value of 281) refers the reader to a page from a bicycle owner's manual to determine how to ensure the seat is in the proper position. The reader must locate the section labelled "Fitting the bicycle" and then identify and summarise the correct information in writing, making sure the conditions stated are contained in the summary.

A second level 3 task, receiving a difficulty value of 310, directs the reader to look at a set of four film reviews to determine which review was least favourable. Some reviews rate films using points or some graphic such as stars; these reviews contain no such indicators. The reader needs to glance at the text of each review to compare what is said in order to judge which film received the worst rating.

Another level 3 question involves an article about cotton diapers. Here readers are asked to write three reasons why the author prefers to use cotton diapers over disposable ones. This task is relatively difficult (318) because of several variables. First, the reader has to provide several answers requiring text-based inferences. Nowhere in the text does the author say, "I prefer cotton diapers because ...". These inferences are made somewhat more difficult because the type of information requested is a "reason" rather than something more concrete such as a date or person. And finally, the text contains information that may distract the reader.

Prose level 4 Score range: 326 to 375

These tasks require readers to perform multiple-feature matching or to provide several responses where the requested information must be identified through text-based inferences. Tasks at this level may also require the reader to integrate or contrast pieces of information, sometimes presented in relatively lengthy texts. Typically, these texts contain more distracting information, and the information requested is more abstract.

One task falling within level 4 (338) directs readers to use the information from a pamphlet about hiring interviews to "write in your own words one difference between the panel interview and the group interview". Here readers are presented with brief descriptions of each type of interview; then, rather than merely locating a fact about each or identifying a similarity, they need to integrate what they have read to infer a characteristic on which the two types of interviews differ. Experience from other large-scale assessments reveals that tasks in which readers are asked to contrast information are more difficult, on average, than tasks in which they are asked to find similarities.

The Hiring Interview

Preinterview

Try to learn more about the business. What products does it manufacture or services does it provide? What methods or procedures does it use? This information can be found in trade directories, chamber of commerce or industrial directories, or at your local employment office.

Find out more about the position. Would you replace someone or is the position newly created? In which departments or shops would you work? Collective agreements describing various standardized positions and duties are available at most local employment offices. You can also contact the appropriate trade union.

The Interview

Ask questions about the position and the business. Answer clearly and accurately all questions put to you. Bring along a note pad as well as your work and training documents.

The Most Common Types of Interview

One-on-one: Self explanatory.

Panel: A number of people ask you questions and then compare notes on your application.

Group: After hearing a presentation with other applicants on the position and duties, you take part in a group discussion.

Postinterview

Note the key points discussed. Compare questions that caused you difficulty with those that allowed you to highlight your strong points. Such a review will help you prepare for future interviews. If you wish, you can talk about it with the placement officer or career counsellor at your local employment office.

Prose level 5 Score range: 376 to 500

Tasks at this level typically require the reader to search for information in dense text that contains a number of plausible distractors. Some require readers to make high-level inferences or to use specialised knowledge.

There is one level 5 task in this assessment, with a difficulty value of 377. Readers are required to look at an announcement from a personnel department and "list two ways in which CIEM (an employee support initiative within a company) helps people who will lose their jobs because of a departmental reorganisation." Responding correctly requires readers to search through this text to locate the embedded sentence "CIEM acts as a mediator for employees who are threatened with dismissal resulting from reorganisation, and assists with finding new positions when necessary." This task is difficult because the announcement is organised around information that is different from what is being requested in the question. Thus, while the correct information is located in a single sentence, this information is embedded under a list of headings describing CIEM's activities for employees looking for other work. This list of headings serves as an excellent set of distractors for the reader who does not search for or locate the phrase containing the conditional information stated in the directive – that is, those who lose their jobs because of a departmental reorganisation.

4.4.2

Document Literacy

Adults often encounter materials such as schedules, charts, graphs, tables, maps and forms at home, at work, or when travelling in their communities. The knowledge and skills needed to process information contained in these documents is therefore an important aspect of literacy in a modern society. Success in processing documents appears to depend at least in part on the ability to locate information in a variety of displays, and to use this information in a number of ways. Sometimes procedural knowledge may be required to transfer information from one source to another, as is necessary in completing applications or order forms.

Thirty-four tasks are ordered along the IALS document literacy scale from 182 to 408, as the result of responses of adults from each of the participating countries. These tasks are distributed as follows: Level 1, 6 tasks; Level 2, 12 tasks; Level 3, 13 tasks; Level 4, 2 tasks; and Level 5, 1 task. By examining tasks associated with these proficiency levels, characteristics that are likely to make particular document tasks more or less difficult can be identified. There are basically four types of questions associated with document tasks: locating, cycling, integrating and generating. Locating tasks require the reader to match one or more features of information stated in the question to either identical or synonymous information given in the document. Cycling tasks require the reader to locate and match one or more features of information, but differ from locating tasks in that they require the reader to engage in a series of feature matches to satisfy conditions given in the question. The integrating tasks typically require the reader to compare and contrast information in adjacent parts of the document. In the generating tasks, readers must produce a written response by processing information found in the document and by making text-based inferences or drawing on their own background knowledge.

As with the prose tasks, each type of question extends over a range of difficulty as a result of combining other variables:

CANCO

CANCO Manufacturing Company
Personnel Department

Centre on Internal and External Mobility

What is CIEM?

CIEM stands for Centre on Internal and External Mobility, an initiative of the personnel department. A number of workers of this department work in CIEM, together with members from other departments and outside career consultants.
CIEM is available to help employees in their search for another job inside or outside the Canco Manufacturing Company.

What does CIEM do?

CIEM supports employees who are seriously considering other work through the following activities:

- ***Job Data Bank***

After an interview with the employee, information is entered into a data bank that tracks job seekers and job openings at Canco and at other manufacturing companies.

- ***Guidance***

The employee's potential is explored through career counselling discussions.

- ***Courses***

Courses are being organized (in collaboration with the department for information and training) that will deal with job search and career planning.

- ***Career Change Projects***

CIEM supports and coordinates projects to help employees prepare for new careers and new perspectives.

- ***Mediation***

CIEM acts as a mediator for employees who are threatened with dismissal resulting from reorganization, and assists with finding new positions when necessary.

How much does CIEM cost?

Payment is determined in consultation with the department where you work. A number of services of CIEM are free. You may also be asked to pay, either in money or in time.

How does CIEM work?

CIEM assists employees who are seriously considering another job within or outside the company.
That process begins by submitting an application. A discussion with a personnel counsellor can also be useful. It is obvious that you should talk with the counsellor first about your wishes and the internal possibilities regarding your career. The counsellor is familiar with your abilities and with developments within your unit.
Contact with CIEM in any case is made via the personnel counsellor. He or she handles the application for you, after which you are invited to a discussion with a CIEM representative.

For more information

The personnel department can give you more information.

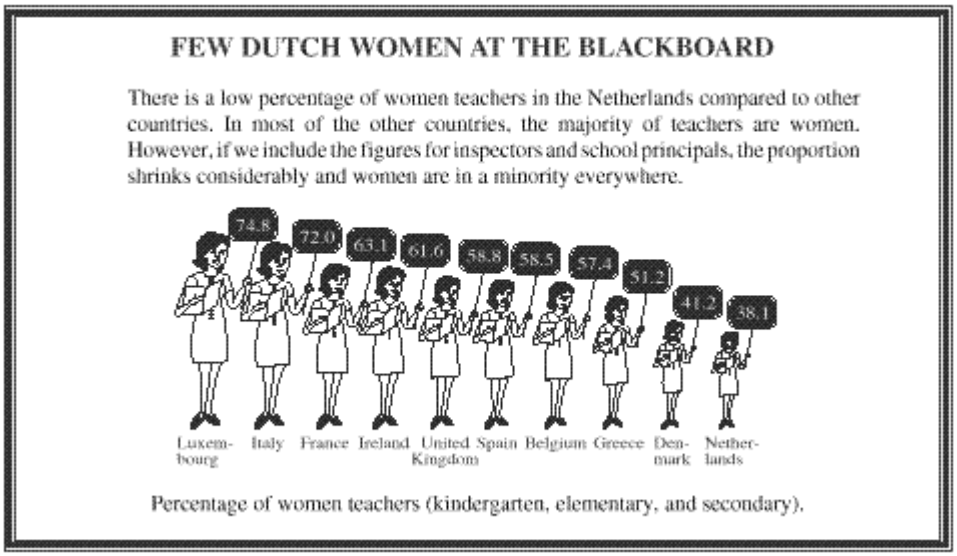
- the number of categories or features of information in the question the reader must process or match;
- the number of categories or features of information in the document that seem plausible or correct because they share some but not all of the information with the correct answer;
- the extent to which the information asked for in the question is clearly related to the information stated in the document;
- the structure and content of the document.

A more detailed discussion of the five levels of document literacy follows.

Document level 1 Score range: 0 to 225

Most of the tasks at this level require the reader to locate a single piece of information based on a literal match. Distracting information, if present, is typically located away from the correct answer. Some tasks may direct the reader to enter personal information onto a form.

One document task at this level (with a difficulty value of 188) directs the reader to identify from a chart the percentage of teachers from Greece who are women. The chart displays the percentages of women teachers from various countries. Only one number appears on the chart for each country.



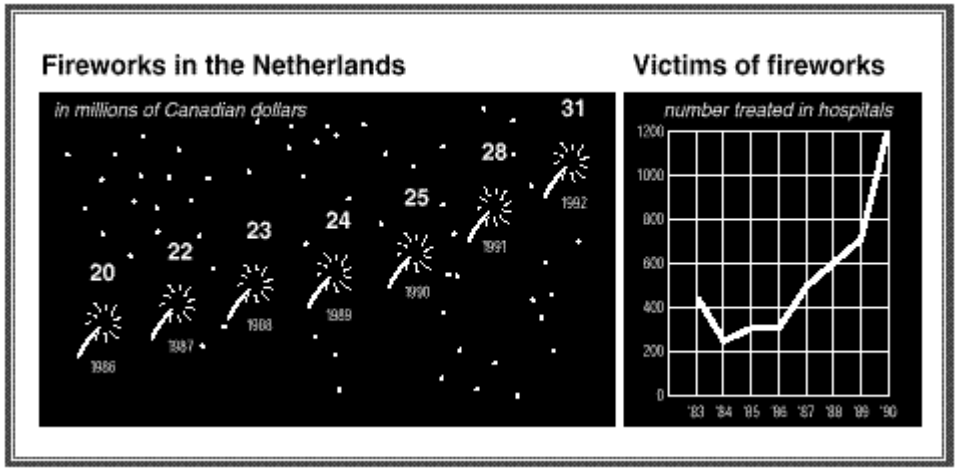
A similar task involves a chart from a newspaper showing the expected amounts of radioactive waste by country. This task, which has a difficulty value of 218, directs the reader to identify the country that is projected to have the smallest amount of waste by the year 2000. Again, there is only one percentage associated with each country; however, the reader must first identify the percentage associated with the smallest amount of waste, and then match it to the country.

Document level 2 Score range: 226 to 275

Document tasks at this level are a bit more varied. While some still require the reader to match a single feature, more distracting information may be present or the match may require a low-level inference. Some tasks at this level may require the reader to enter information onto a form or to cycle through information in a document.

One level 2 task on the document scale (242) directs the reader to look at a chart to identify the year in which the fewest people in the Netherlands were injured by fireworks. Part of what perhaps makes this task somewhat more difficult than those in level 1 is that two charts are presented instead of just one. One, labelled “Fireworks in the Netherlands”, depicts years and numbers representing funds spent in millions of Canadian dollars, whereas the other, “Victims of fireworks”, uses a line to show numbers of people treated in hospitals. It is worth noting that in a second version of the assessment this label was changed to read “number injured.”

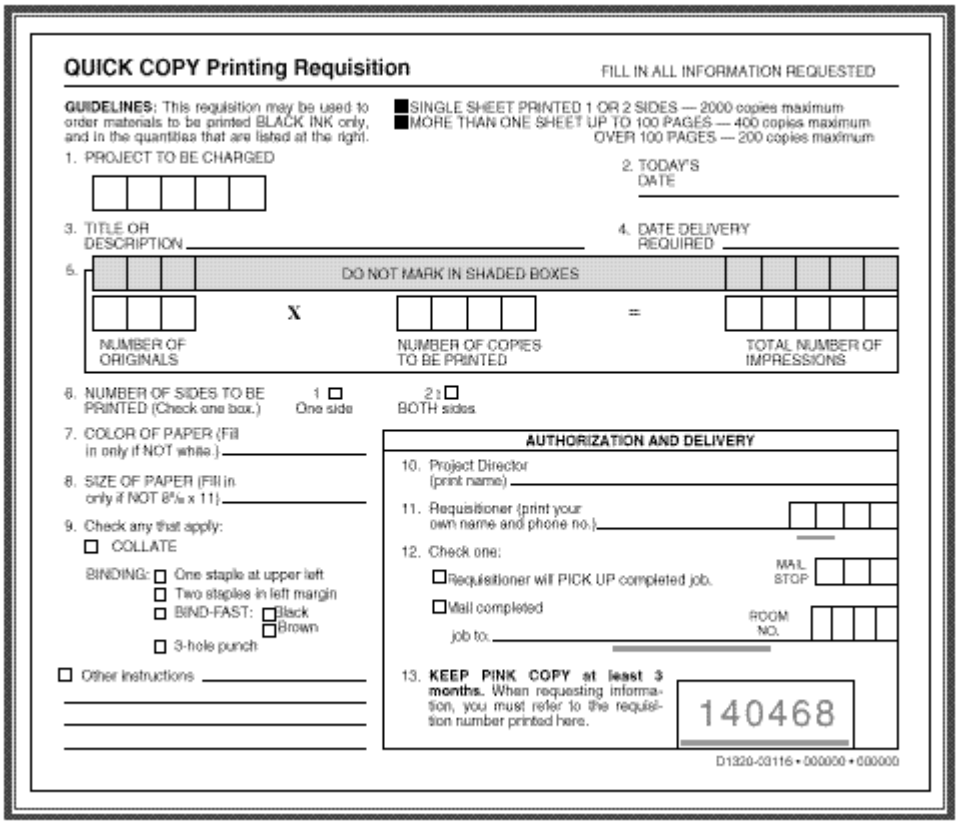
Several other tasks falling within level 2 direct the reader to use information given to complete a form. In one case they are asked to fill out an order form to purchase tickets to see a play on a particular day and at a particular time. In another, readers are asked to complete the availability section of an employment application based on information provided that included: the total number of hours they are willing to work, the hours they are available, how they heard about the job, and the availability of transportation.



Document level 3 Score range: 276 to 325

Tasks at this level are varied. Some require the reader to make literal or synonymous matches, but usually the reader must take conditional information into account or match on the basis of multiple features of information. Some require the reader to integrate information from one or more displays of information. Others ask the reader to cycle through a document to provide multiple responses.

One task falling around the middle of level 3 in difficulty (with a value of 295) involves the fireworks charts shown earlier (see Document level 2). This task directs the reader to write a brief description of the relationship between sales and injuries based on the information shown in the two graphs. A second task, falling at high end of level 3 (321), involves the use of a quick copy printing requisition form that might be found in the workplace. The task asks the reader to state whether or not the quick copy centre would make 300 copies of a statement that is 105 pages long. In responding to this directive, the reader must determine whether conditions stated in the question meet those provided in the requisition form.



QUICK COPY Printing Requisition

FILL IN ALL INFORMATION REQUESTED

GUIDELINES: This requisition may be used to order materials to be printed BLACK INK only, and in the quantities that are listed at the right.

■ SINGLE SHEET PRINTED 1 OR 2 SIDES — 2000 copies maximum
■ MORE THAN ONE SHEET UP TO 100 PAGES — 400 copies maximum
OVER 100 PAGES — 200 copies maximum

1. PROJECT TO BE CHARGED

2. TODAY'S DATE

3. TITLE OR DESCRIPTION

4. DATE DELIVERY REQUIRED

5. DO NOT MARK IN SHADED BOXES

NUMBER OF ORIGINALS x NUMBER OF COPIES TO BE PRINTED = TOTAL NUMBER OF IMPRESSIONS

6. NUMBER OF SIDES TO BE PRINTED (Check one box.) 1 ☐ One side 2 ☐ BOTH sides

7. COLOR OF PAPER (Fill in only if NOT white.)

8. SIZE OF PAPER (Fill in only if NOT 8½ x 11)

9. Check any that apply:
☐ COLLATE
BINDING: ☐ One staple at upper left
☐ Two staples in left margin
☐ BIND-FAST: ☐ Black ☐ Brown
☐ 3-hole punch

☐ Other instructions

AUTHORIZATION AND DELIVERY

10. Project Director (print name)

11. Requisitioner (print your own name and phone no.)

12. Check one:
☐ Requisitioner will PICK UP completed job. MAIL STOP
☐ Mail completed job to: ROOM NO.

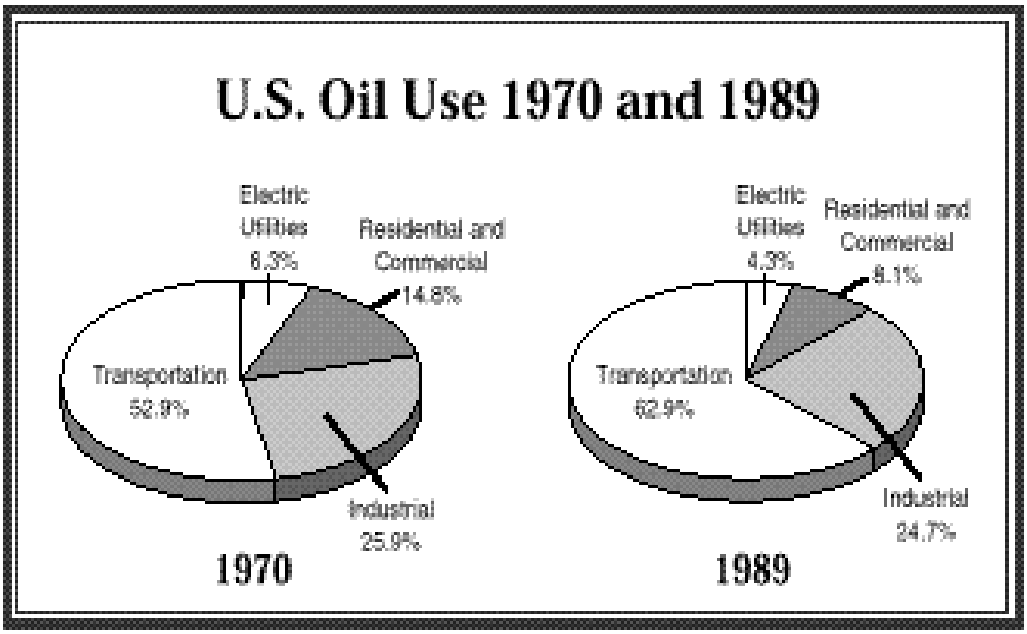
13. **KEEP PINK COPY at least 3 months.** When requesting information, you must refer to the requisition number printed here. 140468

D1320-03116 • 000000 • 000000

Document level 4 Score range: 326 to 375

Tasks at this level, like those at the previous levels, ask the reader to match on the basis of multiple features of information, to cycle through documents, and to integrate information; frequently, however, these tasks require the reader to make higher-order inferences to arrive at the correct answer. Sometimes the document contains conditional information that must be taken into account by the reader.

One of the two tasks falling within this level (341) asks the reader to look at two pie charts showing oil use for 1970 and 1989. The question directs the reader to summarise how the percentages of oil used for different purposes changed over the specified period. Here the reader must cycle through the two charts, comparing and contrasting the percentages for each of the four stated purposes, and then generate a statement that captures these changes.



Document level 5 Score range: 376 to 500

Tasks at this level require the reader to search through complex displays of information that contain multiple distractors, to make high-level inferences, process conditional information, or use specialised knowledge.

The only level 5 task in this international assessment (with a difficulty value of 408) involves a page taken from a consumer magazine rating clock radios. The reader is asked for the average advertised price for the "basic" clock radio receiving the highest overall score. This task requires readers to process two types of conditional information. First, they need to identify the clock radio receiving the highest overall score while distinguishing among the three types reviewed: "full-featured", "basic" and those "with cassette player". Second, they need to locate a price. In making this final match, they need to notice that two are given: the suggested retail price, followed by the average advertised price.

The same document is used for a second and considerably easier task that falls at the low end of level 4 (327). The reader is asked "which full-featured radio is rated the highest on performance". Again, it is necessary to find the correct category of clock radio, but the reader needs to process fewer conditions. All that is required is to distinguish between the rating for "Overall Score" and that for "Performance." It is possible that some adults note the distractor ("Overall Score") rather than the criterion specified in the question, "Performance". Another factor that likely contributes to this task's difficulty is that "Overall Score" is given a numerical value while the other features are rated by a symbol. Also, some adults may find the correct category ("Performance") but select the first radio listed, assuming it performed best. The text accompanying the table indicates that the radios are rated within a category by an overall score; it is easy to imagine that some people may have equated overall score with overall performance.

RATINGS

Clock radios

Listed by types, within types, listed in order of overall score. Differences in score of 4 points or less were not deemed significant.

1 Brand and model. If you can't find a model, call the company. Phone numbers are listed on page 736.

2 Price. The manufacturer's suggested or approximate retail price, followed by the average advertised price.

3 Dimensions. To the nearest centimetre.

4 Overall score. A composite, encompassing all our tests and judgments. A "perfect" radio would have earned 100 points.

5 Convenience. This composite judgment reflects such things as the legibility of the display, the ease of tuning the radio and setting the alarm, and the presence or absence of useful features.

6 Performance. An overall judgment reflecting performance in our tests of: sensitivity and selectivity; tuning ease; capture ratio, the ability to bring in the stronger of two stations on the same frequency; image rejection, the ability to ignore signals from just above the band; resistance to interference from signals bouncing off aircraft and such.

7 Sensitivity. How well each radio received a station with little interference.

8 Selectivity. How well each radio received clearly a weak station next to a strong one on the dial.

9 Tone quality. Based mainly on computer analysis of the speaker's output and on listening tests, using music from CDs. No model produced high-fidelity sound.

10 Reversible time-setting. This useful feature makes setting clock and alarm times easy. If you overshoot the desired setting, you simply back up.

11 Dual alarm. Lets you set two separate wake-up times.

1 Brand and model	Price 2	Dimensions, H×W×D, cm 3	Overall Score 4	Convenience 5	Performance 6	Sensitivity 7	Selectivity 8	Tone quality 9	Reversible time setting 10	Dual alarm 11	Warranty, months	Advantages	Disadvantages	Comments
Full-featured clock radios														
RCA RP-3690	$50/$40	8x25x18	86	●●●●●	✔	✔	12	A,B,D,H,J,L,O,T,U		*A*
Sony ICF-C303	50/45	5x20x15	84	●●●○●	✔	✔	12	C,E,F,I,N,T		*C*
Panasonic RC-X220	50/45	10x28x13	82	●●●●○	✔	✔	12	A,G,K,M,O,S,T,U	b,c	*A*
Realistic 272	50/30	5x28x15	79	●○●●●	✔	✔	3	A,G,H,K,O,T		*D*
Magnavox AJ3900	65/—	15x38x13	78	○●●●●	—	✔	3	D,G,K,M,O,R,T	b,g	*B*
Emerson AK2745	39/20	8x28x15	70	○●●●○	✔	✔	3	G,O	g	*K*
Soundesign 3753	20/20	8x23x13	62	○●●○○	✔	✔	3	J,Q	d,h	*J*
Basic clock radios														
Realistic 263	28/18	10x20x10	74	○●●○●	—	—	3	A,D,H,O,P,U	h	—
Soundesign 3622	12/10	5x20x13	68	●●●●●	—	—	3	U	d	*L*
Panasonic RC-6064	18/15	5x20x13	67	●●●○○	—	—	12	—	b,c	—
General Electric 7-4612	13/10	6x20x13	66	●□●●○	—	—	12	A,D	b,g	—
Lloyds CR001	20/15	5x18x13	64	●○●○●	—	—	3	U	—	—
Sony ICF-C240	15/13	5x18x15	63	●□○○○	—	—	12	—	f,g	—
Emerson AK2720	19/10	5x20x13	61	●○●●○	—	—	3	O,T	e	*K*
Gran Prix D507	15/10	5x18x10	54	●●○●●	—	—	3	—	d	—
Clock radios with cassette player														
General Electric 7-4965	60/50	10x30x15	85	●●●●●	✔	✔	12	A,D,G,H,K,O,S,T	—	*B,E*
Pansonic RC-X250	[1]	10x33x13	76	●●○●●	✔	✔	12	A,G,K,O,R,U	b,c	*A,H*
Sony ICF-CS650	75/65	15x28x15	74	○●●○●	✔	✔	12	G,R,T,U	c,f,i	A,F,H
Soundesign 3844MGY	40/30	13x30x13	62	○●●●●	—	—	3	G,K,J,S,U		*F,G,I,M*

[1] *Discontinued. Replaced by* **RC-X260**, *$79 list and $60 average advertised sale price.*

Features in Common

All: • Permit snooze time of about 8 min. • Retain time settings during short power failures.
Except as noted, all have: • Battery backup for clock and alarm memory. • Red display digits 1 cm. high. • Sleep-time radio play for up to 60 min. before automatic shutoff. • Switch to reset alarm.

Keys to Advantages

A–Alarm works despite power failure.
B–Shows actual time plus up to 2 alarm times.
C–Twin alarms settable for 2 different stations.
D–Tone alarm has adjustable volume control.
E–Memory needs no battery.
F–Digital tuner with presettable stations.
G–Tuner can receive in stereo.
H–Battery-strength indicator.
I–Illuminated tuning dial.
J–Illuminated tuning pointer.
K–Earphone jack.
L–Nap timer.
M–Audio input for tape deck or CD player.
N–Display can show date and time.
O–Display has high/low brightness switch.
P–Display has larger digits than most.
Q–Night light—adjusts for room light.
R–Bass-boost tone control.
S–Treble-cut tone control.
T–Better than most in tuning ease.
U–Better than most in image rejection.

Key to Disadvantages

a–Possible to reset time by accident.
b–Controls for time-setting or dimmer inconveniently located on radio's bottom or rear.
c–Display dimmer than most in brightly lit room.
d–Radio volume must be turned completely down for alarm buzzer to sound.
e–Lacks alarm buzzer; radio is sole alarm.
f–Lacks indication alarm is set.
g–Lacks alarm-reset button.
h–Time-setting lacks fast reverse.
i–No slow forward, fast reverse for time setting.

Key to Comments

A–Display shows green digits.
B–Display shows blue digits.
C–Display uses LCD (liquid crystal) digits.
D–Terminals for external antenna.
E–3-position graphic equalizer.
F–Cassette player lacks Record function.
G–Cassette player lacks Rewind function.
H–Model permits wake-up to cassette play.
I–Cassette-deck flutter worse than most.
J–Warranty repairs cost $3 for handling.
K–Warranty repairs cost $3.50 for handling.
L–Warranty repairs cost $6 for handling.
M–Warranty repairs cost $10 for handling.

4.4.3

Quantitative Literacy

Since adults are frequently required to perform arithmetic operations in everyday life, the ability to perform quantitative tasks is another important aspect of literacy. These skills may at first seem to differ fundamentally from those associated with prose and document literacy, and therefore to extend the concept of literacy beyond its traditional limits. Experience in North America with large-scale assessments of adults indicates that the processing of printed information plays an important role in affecting the difficulty of tasks along the quantitative scale (Montigny et al., 1991; Kirsh et al., 1993).

In general, it appears that many individuals can perform single arithmetic operations when both the numbers and operations are made explicit. However, when the numbers to be used must be located in and extracted from different types of documents that contain other similar but irrelevant information, when the operations to be used must be inferred from printed directions, and when multiple operations must be performed, the tasks become increasingly difficult.

The IALS quantitative literacy scale contains 33 tasks ranging from 229 to 408 in difficulty. These tasks are distributed as follows: Level 1, 1 task; Level 2, 9 tasks; Level 3, 16 tasks; Level 4, 5 tasks; and Level 5, 2 tasks. The difficulty of these tasks – and therefore, their placement along the scale – appears to be a function of several factors including:

- the particular arithmetic operation the task requires;
- the number of operations needed to perform the task successfully;
- the extent to which the numbers are embedded in printed materials;
- the extent to which an inference must be made to identify the type of operation to be performed.

The five levels of quantitative literacy are described in detail below.

Quantitative level 1 Score range: 0 to 225

Although no quantitative tasks used in the assessment fall below the score value of 225, experience suggests that such tasks would require the reader to perform a single, relatively simple operation (usually addition) for which either the numbers are clearly noted in the given document and the operation is stipulated, or the numbers are provided and the operation does not require the reader to find the numbers.

The easiest quantitative task (225) directs the reader to complete an order form. The last line on this form says “Total with Handling”. The line above it says “Handling Charge $2.00”. The reader simply has to add the $2.00 to the $50.00 entered on a previous line to indicate the cost of the tickets. In this task, one of the numbers is stipulated; the operation is easily identified from the word “total”; and the operation does not require the reader to perform the “borrow” or “carry-over” function of addition. Moreover, the form itself features a simple column format, further facilitating the task for the reader.

Quantitative level 2 Score range: 226 to 275

Tasks at this level typically require readers to perform a single arithmetic operation (frequently addition or subtraction), using numbers that are easily located in the text or document. The operation to be performed may be easily inferred from the wording of the question or the format of the material (for example, a bank deposit or order form).

A typical level 2 task on the quantitative scale directs the reader to use a weather chart in a newspaper to determine how many degrees warmer today's high temperature is expected to be in Bangkok than in Seoul. Here the reader must cycle through the table to locate the two temperatures and then subtract one from the other to determine the difference. This task received a difficulty value of 255.

A similar but slightly more difficult task (268) requires the reader to use the chart about women in the teaching profession that is displayed in level 1 for the document scale. This task directs the reader to calculate the percentage of men in the teaching profession in Italy. Both this task and the one just mentioned involve calculating the difference between two numbers. In the former, however, both temperatures could be identified in the table from the newspaper. For the task involving male teachers in Italy, the reader needs to make the inference that the percentage is equal to 100 per cent minus the percentage of female teachers.

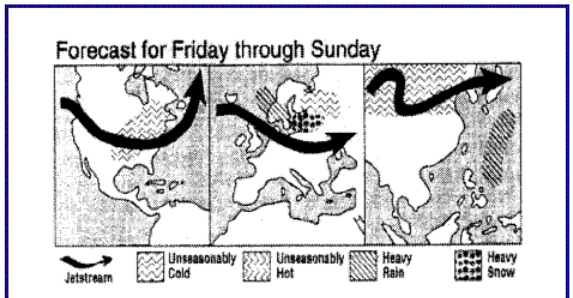
WEATHER

Europe

	Today High C	Today Low C	Today W	Tomorrow High C	Tomorrow Low C	Tomorrow W
Algarve	19	7	s	21	9	s
Amsterdam	11	6	pc	12	7	pc
Ankara	17	7	pc	19	8	pc
Athens	22	15	pc	23	14	pc
Barcelona	16	8	s	14	9	s
Belgrade	14	6	pc	10	1	c
Berlin	8	2	c	6	1	c
Brussels	11	6	pc	14	7	pc
Budapest	9	1	pc	9	2	c
Copenhagen	7	1	r	6	2	c
Costa del Sol	21	8	s	21	10	s
Dublin	10	6	pc	13	8	pc
Edinburgh	10	6	c	10	6	c
Florence	11	5	s	14	6	s
Frankfurt	12	6	pc	13	4	pc
Geneva	9	2	s	12	4	s
Helsinki	-1	-7	sf	-3	-10	pc
Istanbul	17	10	pc	15	9	sh
Las Palmas	26	18	pc	27	18	pc
Lisbon	19	9	s	19	10	s
London	12	5	pc	13	7	pc
Madrid	17	3	s	18	4	s
Milan	9	3	s	13	6	s
Moscow	1	-3	r	-3	-11	sf
Munich	11	3	pc	12	6	pc
Nice	14	7	s	15	8	s
Oslo	4	-4	c	5	-2	c
Paris	12	6	pc	13	6	pc
Prague	11	1	pc	8	2	c
Reykjavik	4	2	r	6	-1	c
Rome	20	12	s	20	10	s
St. Petersburg	-1	-7	sf	-4	-12	pc
Stockholm	1	-5	sn	-2	-7	c
Strasbourg	12	5	pc	15	7	pc
Tallinn	-1	-7	sf	-4	-10	pc
Venice	10	3	s	11	4	s
Vienna	9	-1	pc	10	2	c
Warsaw	8	2	sh	6	1	c
Zurich	8	0	s	9	1	pc

Oceania

Auckland	20	14	s	17	11	sh
Sydney	27	17	pc	25	16	pc



North America
Cold weather will engulf the Midwestern and Northeastern United States Friday and over the weekend. Although it will be cold in Chicago, Toronto and New York City, the weather is expected to be dry. Los Angeles will have some sunshine and seasonable temperatures each day.

Europe
Western and central Europe will have a spell of mild weather Friday into the weekend. London and Paris will have dry weather with some sunshine Friday into Sunday. Rain will continue to soak southwestern Norway. Snow will blanket the area from Minsk to Moscow.

Asia
Typhoon Elsie will probably stay to the east of the Philippines and south of Japan Friday and the weekend. Some rain is apt to fall in Seoul and there could even be a little ice or snow. Cold air will pour into Beijing and snow is a possibility. Hong Kong will start the weekend warm.

Middle East

	Today High C	Today Low C	Today W	Tomorrow High C	Tomorrow Low C	Tomorrow w
Beirut	28	19	pc	29	20	s
Cairo	29	20	pc	28	19	pc
Damascus	24	12	s	26	14	s
Jerusalem	27	15	s	26	14	s
Riyadh	34	13	s	32	13	s

Latin America

	Today High C	Today Low C	Today W	Tomorrow High C	Tomorrow Low C	Tomorrow W
Buenos Aires	23	11	pc	26	13	s
Caracas	29	20	s	31	18	s
Lima	23	17	c	23	16	c
Mexico City	23	11	sh	23	12	pc
Rio de Janiero	32	22	s	28	21	sh
Santiago	24	4	s	22	6	pc

Legend: s-sunny, pc-partly cloudy, c-cloudy, sh-showers, t-thunderstorms, r-rain, sf-snow flurries, sn-snow, i-ice, W-Weather. **All maps, forecasts and data provided by Accu-Weather, Inc. © 1992**

Asia

	Today High C	Today Low C	Today W	Tomorrow High C	Tomorrow Low C	Tomorrow W
Bangkok	32	22	pc	30	23	s
Beijing	11	0	s	8	2	pc
Hong Kong	30	23	s	29	22	pc
Manila	31	25	s	31	25	sh
New Delhi	31	13	s	32	16	s
Seoul	14	6	pc	14	4	pc
Shanghai	22	10	pc	24	12	s
Singapore	31	24	pc	28	23	sh
Taipei	26	21	pc	26	19	pc
Tokyo	18	9	pc	17	7	pc

Africa

Algiers	27	14	s	26	13	s
Cape Town	20	11	sh	18	11	pc
Casablanca	20	14	c	21	11	pc
Harare	34	17	s	32	18	pc
Lagos	30	24	pc	29	24	pc
Nairobi	27	12	pc	26	13	pc
Tunis	27	17	pc	17	14	pc

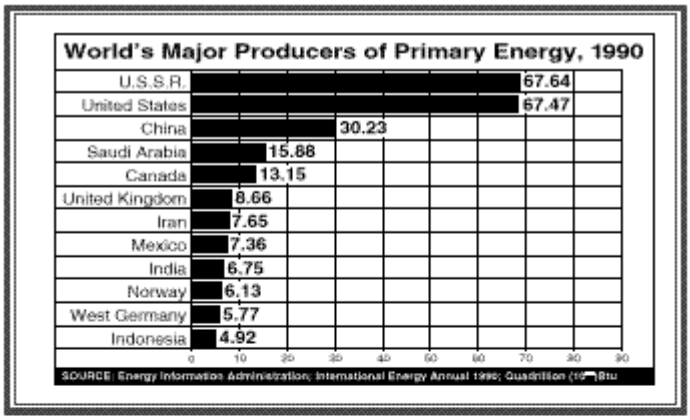
North America

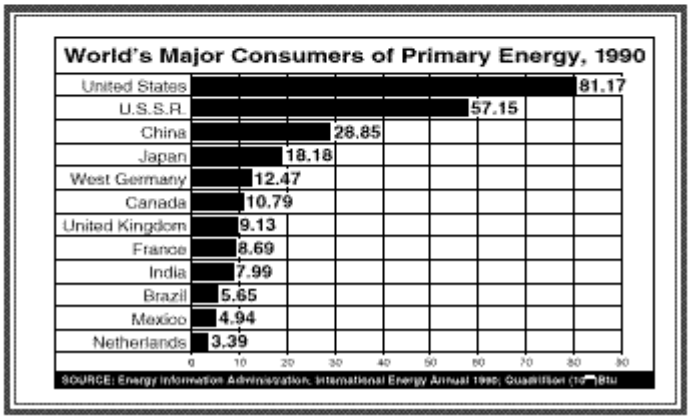
Anchorage	0	-2	c	3	0	sh
Atlanta	14	4	pc	8	2	pc
Boston	15	4	c	8	-1	pc
Chicago	2	-5	c	-2	-8	pc
Denver	8	-3	pc	4	-6	sn
Detroit	4	-2	c	4	-5	pc
Honolulu	31	20	s	31	21	pc
Houston	15	3	pc	12	6	pc
Los Angeles	28	14	s	24	13	s
Miami	30	22	pc	29	21	pc
Minneapolis	-1	-8	c	1	-7	pc
Montreal	7	-2	sf	4	-3	c
Nassau	31	22	pc	28	21	sh
New York	14	4	r	10	2	pc
Phoenix	23	11	pc	22	8	s
San Fran.	20	11	pc	21	8	s
Seattle	11	6	pc	13	7	r
Toronto	6	-3	c	3	-3	c
Washington	14	6	r	11	4	pc

Quantitative level 3 Score range: 276 to 325

Tasks at this level typically require the reader to perform a single operation. However, the operations become more varied – some multiplication and division tasks are included. Sometimes the reader needs to identify two or more numbers from various places in the document, and the numbers are frequently embedded in complex displays. While semantic relation terms such as "how many" or "calculate the difference" are often used, some of the tasks require the reader to make higher-order inferences to determine the appropriate operation.

One task located at 302 on the quantitative scale directs the reader to look at two graphs containing information about consumers and producers of primary energy. The reader is asked to calculate how much more energy Canada produces than it consumes. Here the operation is not facilitated by the format of the document; the reader must locate the information using both bar graphs. Another task involving this document directs the reader to calculate the total amount of energy in quadrillion (10^{15}) BTU (British Thermal Unit) consumed by Canada, Mexico and the United States. This task, which falls at 300 on the scale, requires the reader to add three numbers. Presenting two graphs likely increases the difficulty; some respondents may perform the appropriate calculation for the three countries specified using the producer energy chart rather than the consumer energy chart.





Another task at this level involves the fireworks chart shown previously for the document scale. The reader is asked to calculate how many more people were injured in 1989 than in 1988. What contributes to this task receiving a difficulty value of 293 is that one of the numbers is not given in the line graph; the reader needs to interpolate the number from information provided along the vertical axis.

A task located at 280 on the scale asks readers to look at a recipe for scrambled eggs with tomatoes. The recipe gives the ingredients for four servings: 3 tablespoons of oil, 1 garlic clove, 1 teaspoon of sugar, 500 grams of fresh red tomatoes and 6 eggs. They are then asked to determine the number of eggs they will need if they are using the recipe for six people. Here they must know how to calculate or determine the ratio needed. This task is somewhat easier than might be expected given others at the same level, perhaps because people are familiar with recipes and with manipulating them to fit a particular situation.

Another question using this recipe asks the reader to determine the amount of oil that would be needed if the recipe were being used for two people. This task received a value of 253 on the scale; a larger percentage of respondents found it easier to halve an ingredient than to increase one by 50 per cent. It is not clear why this is so. It may be that some of the respondents have an algorithm for responding to certain familiar tasks that does not require them to apply general arithmetic principles.

Quantitative level 4 Score range: 326 to 375

With one exception, the tasks at this level require the reader to perform a single arithmetic operation where typically either the quantities or the operation are not easily determined. That is, for most of the tasks at this level, the question or directive does not provide a semantic relation term such as “how many” or “calculate the difference” to help the reader.

One task at this level involves a compound interest table. It directs the reader to “calculate the total amount of money you will have if you invest $100 at a rate of 6 per cent for 10 years.” This task received a difficulty value of 348, in part because many people treated this as a document rather than a quantitative task and simply looked up the amount of interest that would be earned. They likely forgot to add the interest to their $100 investment.

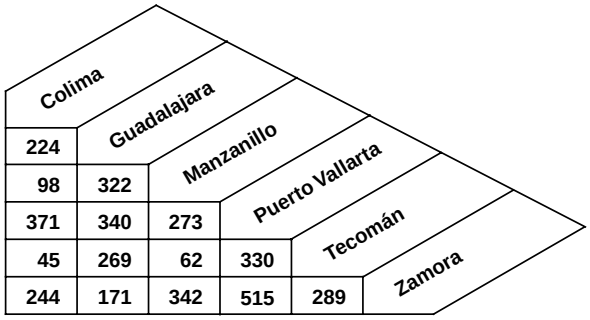
Compound Interest
Compounded Annually

Principal	Period	4%	5%	6%	7%	8%	9%	10%	12%	14%	16%
$100	1 day	0.011	0.014	0.016	0.019	0.022	0.025	0.027	0.033	0.038	0.044
	1 week	0.077	0.096	0.115	0.134	0.153	0.173	0.192	0.230	0.268	0.307
	6 mos	2.00	2.50	3.00	3.50	4.00	4.50	5.00	6.00	7.00	8.00
	1 year	4.00	5.00	6.00	7.00	8.00	9.00	10.00	12.00	14.00	16.00
	2 years	8.16	10.25	12.36	14.49	16.64	18.81	21.00	25.44	29.96	34.56
	3 years	12.49	15.76	19.10	22.50	25.97	29.50	33.10	40.49	48.15	56.09
	4 years	16.99	21.55	26.25	31.08	36.05	41.16	46.41	57.35	68.90	81.06
	5 years	21.67	27.63	33.82	40.26	46.93	53.86	61.05	76.23	92.54	110.03
	6 years	26.53	34.01	41.85	50.07	58.69	67.71	77.16	97.38	119.50	143.64
	7 years	31.59	40.71	50.36	60.58	71.38	82.80	94.87	121.07	150.23	182.62
	8 years	36.86	47.75	59.38	71.82	85.09	99.26	114.36	147.60	185.26	227.84
	9 years	42.33	55.13	68.95	83.85	99.90	117.19	135.79	177.31	225.19	280.30
	10 years	48.02	62.89	79.08	96.72	115.89	136.74	159.37	210.58	270.72	341.14
	12 years	60.10	79.59	101.22	125.22	151.82	181.27	213.84	289.60	381.79	493.60
	15 years	80.09	107.89	139.66	175.90	217.22	264.25	317.72	447.36	613.79	826.55
	20 years	119.11	165.33	220.71	286.97	366.10	460.44	572.75	864.63	1,274.35	1,846.06

Another task at this level requires respondents to read a newspaper article describing a research finding linking allergies to a particular genetic mutation. The question directs the reader to calculate the number of people studied who were found to have the mutant gene. To answer the question correctly, readers must know how to convert the phrase “64 per cent” to a decimal number and then multiply it by the number of patients studied (400). The text provides no clues on how to tackle this problem.

A third task involves a distance chart. Readers are asked to “calculate the total number of kilometres travelled in a trip from Guadalajara to Tecomán and then to Zamora”. Here a semantic relation term is provided, but the format is difficult and the quantities are not easily identified. As a result, this task received a difficulty value of 335. In a level 3 task using the same chart, respondents are asked to determine how much less the distance from Guadalajara to Tecomán is than the distance from Guadalajara to Puerto Vallarta. In that task (308), the quantities are relatively easy to locate.

TABLE OF APPROXIMATE DISTANCES (in kilometres)



Quantitative level 5 Score range: 376 to 500

These tasks require readers to perform multiple operations sequentially, and they must locate features of the problem embedded in the material or rely on background knowledge to determine the quantities or operations needed.

One of the most difficult tasks on the quantitative scale (381) requires readers to look at a table providing nutritional analysis of food and then, using the information given, determine the percentage of calories in a Big Mac® that comes from total fat. To answer this question, readers must first recognise that the information about total fat provided is given in grams. In the question, they are told that a gram of fat has 9 calories. Therefore, they must convert the number of fat grams to calories. Then, they need to calculate this number of calories as a percentage of the total calories given for a Big Mac®. Only one other item on this scale received a higher score.

Nutritional Analysis

	Serving Size	Calories	Protein (g)	Carbohydrates (g)	Total Fat (g)	Saturated Fat (g)	Monounsaturated Fat (g)	Polyunsaturated Fat (g)	Cholesterol (mg)	Sodium (mg)
Sandwiches										
Hamburger	102 g	255	12	30	9	5	1	3	37	490
Cheeseburger	116 g	305	15	30	13	7	1	5	50	725
Quarter Pounder®	166 g	410	23	34	20	11	1	8	85	645
Quarter Pounder® w/Cheese	194 g	510	28	34	28	16	1	11	115	1110
McLean Deluxe™	206 g	320	22	35	10	5	1	4	60	670
McLean Deluxe™ w/Cheese	219 g	370	24	35	14	8	1	5	75	890
Big Mac®	215 g	500	25	42	26	16	1	9	100	890
Filet-O-Fish®	141 g	370	14	38	18	8	6	4	50	730
McChicken®	187 g	415	19	39	19	9	7	4	50	830
French Fries										
Small French Fries	68 g	220	3	26	12	8	1	2.5	0	110
Medium French Fries	97 g	320	4	36	17	12	1.5	3.5	0	150
Large French Fries	122 g	400	6	46	22	15	2	5	0	200
Salads										
Chef Salad	265 g	170	17	8	9	4	1	4	111	400
Garden Salad	189 g	50	4	6	2	1	0.4	0.6	65	70
Chunky Chicken Salad	255 g	150	25	7	4	2	1	1	78	230
Side Salad	106 g	30	2	4	1	0.5	0.2	0.3	33	35
Croutons	11 g	50	1	7	2	1.3	0.1	0.5	0	140
Bacon Bits	3 g	15	1	0	1	0.3	0.2	0.5	1	95

Soft Drinks

	Coca-Cola Classic®				diet Coke®				Sprite®			
	Small	Medium	Large	Jumbo	Small	Medium	Large	Jumbo	Small	Medium	Large	Jumbo
Calories	140	190	260	380	1	1	2	3	140	190	260	380
Carbohydrates (g)	38	50	70	101	0.3	0.4	0.5	0.6	36	48	66	96
Sodium (mg)	15	20	25	40	30	40	60	80	15	20	25	40

4.5

Estimating Literacy Performance Across the Levels

The literacy levels not only provide a means for exploring the progression of information-processing demands across each of the scales, but also can be used to help explain how the proficiencies individuals demonstrate reflect the likelihood they will respond correctly to the broad range of tasks used in this assessment as well as to any task that has the same characteristics. In practical terms, this means that individuals performing at 250 on each scale are expected to be able to perform the average level 1 and 2 tasks with a high degree of proficiency – i.e. with an average probability of a correct response at 80 per cent or higher. It does not mean that they will not be able to perform tasks in levels 3 or higher. They would be expected to do so some of the time, but not consistently.

The three charts given in Tables 4.1 to 4.3 display the probability that individuals performing at selected points on each of the scales will give a correct response to tasks of varying difficulty. For example, a reader whose prose proficiency is 150 has less than a 50 per cent chance of giving a correct response to the level 1 tasks. Individuals whose proficiency score is 200, in contrast, have about an 80 per cent probability of responding correctly to these tasks.

In terms of task demands, it can be inferred that adults performing at 200 on the prose scale are likely to be able to locate a single piece of information in a brief text when there is no distracting information, or if plausible but incorrect information is present but located away from the correct answer. However, these individuals are likely to encounter far more difficulty with tasks in levels 2 through 5. For example, they would have only a 40 per cent chance of performing the average level 2 task correctly, an 18 per cent chance of success with tasks in level 3, and no more than a 7 per cent chance with tasks in levels 4 and 5.

In contrast, respondents demonstrating a proficiency of 300 on the prose scale have about an 80 per cent chance or higher of succeeding with tasks in levels 1, 2 and 3. This means that they demonstrate success with tasks that require them to make low-level inferences and with those that entail taking some conditional information into account. They can also integrate or compare and contrast information that is easily identified in the text. On the other hand, they are likely to encounter difficulty with tasks where they must make more sophisticated text-based inferences, or where they need to process more abstract types of information. These more difficult tasks may also require them to draw on less familiar or more specialised types of knowledge beyond that given in the text. On average, they have about a 50 per cent probability of performing level 4 tasks correctly; with level 5 tasks, their likelihood of responding correctly decreases to 40 per cent.

Similar kinds of interpretations can be made using the information presented for the document and quantitative literacy scales. For example, someone who is at 200 on the quantitative scale has, on average, a 67 per cent chance of responding correctly to level 1 tasks. His or her likelihood of responding correctly decreases to 47 per cent for level 2 tasks, 21 per cent for level 3 tasks, 6 per cent for level 4 tasks and a mere 2 per cent for level 5 tasks. Similarly, readers with a proficiency of 300 on the quantitative scale would have a probability of 92 per cent or higher of responding correctly to tasks in levels 1 and 2. Their average probability would decrease to 81 per cent for level 3 tasks, 57 per cent for level 4 and 20 per cent for level 5.

Table 4.1
Average probabilities of successful performance, prose scale

	Selected proficiency scores				
Prose level	**150**	**200**	**250**	**300**	**350**
			%		
1	48	81	95	99	100
2	14	40	76	94	99
3	6	18	46	78	93
4	2	7	21	50	80
5*	2	6	18	40	68

* Based on one task
Source: Adult Literacy Survey (1994).

Table 4.2
Average probabilities of successful performance, document scale

	Selected proficiency scores				
Document Level	**150**	**200**	**250**	**300**	**350**
			%		
1	40	72	94	99	100
2	20	51	82	95	99
3	7	21	50	80	94
4	4	13	34	64	85
5*	<1	1	3	13	41

* Based on one task
Source: Adult Literacy Survey (1994).

Table 4.3
Average probabilities of successful performance, quantitative scale

	Selected proficiency scores				
Quantitative level	**150**	**200**	**250**	**300**	**350**
			%		
1	34	67	89	97	99
2	21	47	76	92	98
3	7	21	51	81	94
4	1	6	22	57	86
5*	1	2	7	20	53

* Based on one task
Source: Adult Literacy Survey (1994).

4.6

Estimating the Variability of Literacy Tasks Across the Participating Countries

One of the goals in conducting international surveys is to be able to compare populations on common scales. In this study, three literacy scales were used to compare both the distributions of literacy skills and the relationships between literacy skills and a variety of social, educational and labour market variables. The literacy tasks received item parameters that define its difficulty and how well it discriminates among populations of adults. These parameters were determined on the basis of how adults within and across participating countries responded to each task.

Under standard assumptions of IRT, item parameters are thought to be invariant among respondents and among countries as well as subgroups within countries. However, it has been discovered through performing large-scale assessments that this assumption is not always true. Yamamoto (1997) notes that some language/country populations do respond differently to a subset of literacy tasks. As described in the IALS Technical Report (Murray *et al.,* 1997), individual items were dropped from the assessment if at least seven of the original ten language or country populations were shown not to have

the same item parameters – *i.e.* if the response data for a particular item proved to have a poor fit to the item parameters common to the rest of the language or country populations. In addition, if there were items in which only one, two or three countries varied, these countries were allowed to have unique parameters for that item. This resulted in a total of 13 items being dropped from the assessment, with 31 items getting a unique parameter for one language or country population, 16 for two language or country populations, and 6 for three language or country populations. Another way to look at this is that there were a total of 1,010 constraints (114 items minus the 13 dropped times 10 language samples). Of these, unique item parameters were required or allowed in 81 instances, meaning that 92 per cent of the constraints support a common scale across the ten original language or country populations.

These discrepancies were due largely to differences in translations among countries, or to differences in interpretation of scoring rubrics for individual items. The different performance on some items also reflected the variation in language and culture, although no obvious or specific reason could be identified. The fact that not all items had identical item parameters resulted in two types of variation. First, differences could influence the distribution of proficiency scores for a particular language or country group, if only slightly. Analyses indicated that the consequence of using a partially different set of item parameters on the proficiency distribution for a particular population was minimal. For any population, when the proficiency distribution was estimated based either on a set of items which included those common across countries as well as those unique to a given country, or on a set of items which were optimal for a different population, the means and standard deviations of estimated proficiencies differed by less than half of a standard error. Typically, standard errors of estimation ranged between 1 and 3 points on the 500-point scales depending on a particular language or country population.

The second type of variation which results from having a small set of items with unique parameters occurs in the placement of particular tasks along the scales according to their response probability of 80 per cent (RP80). At the beginning of this annex, it was mentioned that a criterion of 80 per cent was used, meaning that tasks were placed along a scale based on the probability that someone with that level of proficiency would have an 80 per cent chance of getting that task and others like it correct. The fact that small subsets of tasks have unique parameters for particular country/language groups results in some tasks falling at different points along each scale.

To evaluate the variability of average probabilities of correct responses (RP80s) for each language or country population, the deviation of RP80s against the common RP80 was examined. It is important to note that no country received all common item parameters. That is, at least one item for each country received a unique set of parameters. However, at least seven of the original language or country populations received common parameters for each of the 101 items. In total, there are 24 language or country groups for which data are currently available to estimate this variation. Nine of the groups are from the first assessment reported in 1995, six are from the assessment cycle reported in 1997 and nine are from the final round of surveys. There were a total of 101 literacy tasks in the assessment so there could be as many as 1,515 deviations (101 times 15).

The mean deviation among the RP80s was 4.7, with a standard deviation of 15.3. This means that the average variation among the RP80s for the literacy tasks was 4.7 points on a 500-point scale, or less than 10 per cent of the 50 points making up a particular literacy level. In addition, a small number of items had large deviations, accounting for a significant percentage of this variation. Only 2 per cent of the deviations observed account for about 35 per cent of the average deviation. In other words, 98 per cent of the deviations have a mean of 3.0, or a 35 per cent reduction from the average of 4.7.

Table 4.4 shows the average deviation of the RP80s for each of the 24 country or language groups; the average is seen to range from a low of 1.1 for the French-speaking Swiss to 10.3 for Hungary.

Table 4.4
Average deviation of RP80 values by country or language group

Australia	7.6	**Germany**	5.3	**Norway (Bokmål)**	2.7
Belgium (Flanders)	5.8	**Great Britain**	5.2	**Poland**	5.4
Canada (English)	3.6	**Hungary**	10.3	**Slovenia**	5.3
Canada (French)	3.2	**Ireland**	4.5	**Sweden**	5.2
Chile	3.5	**Italy**	3.8	**Switzerland (French)**	1.1
Czech Republic	3.7	**Netherlands**	3.4	**Switzerland (German)**	4.0
Denmark	3.2	**New Zealand**	7.2	**Switzerland (Italian)**	6.0
Finland	3.6	**Northern Ireland**	6.9	**United States**	2.0

4.7 Conclusion

One of the goals of large-scale surveys is to provide information that can help policy makers during the decision-making process. Presenting that information in a way that will enhance understanding of what has been measured and the conclusions to be drawn from the data is important to reaching this goal. This guide has offered a framework for understanding the consistency of task responses demonstrated by adults from a number of countries. The framework identifies a set of variables that have been shown to underlie successful performance on a broad array of literacy tasks. Collectively, they provide a means for moving away from interpreting survey results in terms of discrete tasks or a single number, and towards identifying levels of performance sufficiently generalised to have validity across assessments and groups.

The concept of test design is evolving. Frameworks such as the one presented here can assist in that evolution. No longer should testing stop at assigning a numerical value; it should assign meaning to that number. And, as concern ceases to centre on discrete behaviours or isolated observations and focus is more on providing a *meaningful* score, a higher level of measurement is reached (Messick, 1989).

References

KIRSCH, I.S. and MOSENTHAL, P. (1993), "Interpreting the IEA reading literacy scales", in M. Binkley, K. Rust and M. Winglee (Eds.), *Methodological Issues in Comparative Educational Studies: The Case of the IEA Reading Literacy Study,* National Center for Education Statistics, United States Department of Education, Washington, DC.

KIRSCH, I.S., JUNGEBLUT, A., JENKINS, L., and KOLSTAD, A. (Eds.) (1993), *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey,* National Center for Education Statistics, United States Department of Education, Washington, DC.

MESSICK, S. (1989), "Validity", in R. Linn (Ed.), *Educational Measurement*, 3rd edition, Macmillan, New York.

MONTIGNY, G., KELLY, K., and JONES, S. (1991), *Adult Literacy in Canada: Results of a National Study,* Statistics Canada, Catalogue No. 89-525-XPE, Minister of Industry, Science and Technology, Ottawa.

MURRAY, T.S., KIRSCH, I.S., and JENKINS, L. (Eds.) (1997), *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey,* National Center for Education Statistics, United States Department of Education, Washington, DC.

YAMAMOTO, K. (1997), "Scaling and scale linking", in T.S. Murray, I.S. Kirsch, and L. Jenkins (Eds.), *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey,* National Center for Education Statistics, United States Department of Education, Washington, DC.

5.0 Survey Methodology

The survey methodology section details the methodology used in each of the 20 IALS countries in terms of target population, frame coverage and sample design.

5.1

Target Population and Frame Coverage

Each country designed a sample that had to be representative of their civilian, non-institutionalised population aged 16-65. Only a small number of exclusions were acceptable. Tables 5.1a and 5.1b show the high rate of coverage achieved by each of the participating countries.

Countries were encouraged to field sample sizes large enough to yield 3,000 completed cases after non-response, so that secondary analysis and estimates of literacy profiles could be obtained reliably.

TABLE 5.1a

SURVEY COVERAGE AND EXCLUSIONS (FIRST CYCLE)

Country	Coverage (per cent)	Exclusions
Belgium[1] (Flanders)	99	Residents of institutions
Canada	98	Residents of institutions, persons living on Indian reserves, members of the armed forces, residents of the Yukon and Northwest Territories
Germany	98	Residents of institutions
Ireland	100	None
Netherlands	99	Residents of institutions
New Zealand	99	Residents of institutions; offshore islands, onshore islands, waterways and inlets
Poland	99	Persons residing in the country for less than three months
Sweden	98	Persons living in institutions (including those doing their military service), persons living abroad during the survey period
Switzerland (French and German)	99	Residents of institutions; persons without telephones
United Kingdom	97	Residents of institutions; the Scottish Highlands and islands north of the Caledonian Canal
United States	97	Members of the armed forces on active duty, those who reside outside the country, those without a fixed household address

1. The Belgium IALS-sample is representative of the "Flemish Region", excluding Brussels. Therefore, the label "Flanders" is used throughout this publication, rather than the more conventional "Flemish Community".

TABLE 5.1b

SURVEY COVERAGE AND EXCLUSIONS (SECOND CYCLE)

Country	Coverage (per cent)	Exclusions
Chile	98	Residents of institutions; persons in remote areas
Czech Republic	98	Residents of institutions; members of the armed forces; citizens living abroad
Denmark	99	Residents of institutions
Finland	94	Residents of institutions; citizens living abroad; persons with Swedish as the mother tongue
Hungary	99	Residents of institutions; persons without a fixed address
Italy	91	Residents of institutions; persons without telephone; persons with unlisted telephone numbers
Norway	99	Residents of institutions for more than six months
Slovenia	98	Residents of institutions; refugees; foreigners
Switzerland (Italian)	99	Residents of institutions; persons without telephones

The tests could be conducted in more than one language in a country. Canada, Norway and Switzerland chose to do so. In Canada respondents were given a choice of English or French. In Norway surveys were conducted in Bokmål and Nynorsk, but this document presents the Bokmål results only. In Switzerland, samples drawn from French, German or Italian-speaking cantons (mostly Ticino) were required to respond in those respective languages (Rhaeto-Romanic-speaking regions were excluded). For some countries or languages, the target sample size was fixed at 1,500 cases – the minimum number necessary for producing reliable literacy profiles – instead of 3,000 cases.

In all countries, when respondents could not speak the designated test language, attempts were made to collect information through the background questionnaire so as to allow for the imputation of missing literacy information and hence reduce the possibility of biased results.

Tables 5.1c and 5.1d show the target populations and the test languages used in each country. Although the common target population was people aged 16-65, individual countries were free to sample younger or older adults. Canada, Sweden and Switzerland sampled persons at least 16 years of age but with no upper limit, while the Netherlands sampled persons aged 16 to 74, and Australia sampled those aged 15 to 74. Chile also took this opportunity, including young adults 15 years of age.

The total number of respondents aged 16-65 in IALS over both of the two cycles is given in Table 5.1e.

TABLE 5.1c

TEST LANGUAGE, TARGET POPULATION SIZE AND NUMBER OF SURVEY RESPONDENTS (FIRST CYCLE)

Country	Test language	Population aged 16-65	Survey respondents aged 16-65[1]
Belgium (Flanders)	Dutch	4,500,000	2,261
Canada	English	13,700,000	3,130
	French	4,800,000	1,370
Germany	German	53,800,000	2,062
Ireland	English	2,200,000	2,423
Netherlands	Dutch	10,500,000	2,837
New Zealand	English	2,100,000	4,223
Poland	Polish	24,500,000	3,000
Sweden	Swedish	5,400,000	2,645
Switzerland	French	1,000,000	1,433
	German	3,000,000	1,393
United Kingdom	English	37,000,000	6,718
United States	English	161,100,000	3,038

1. Includes records where age was not stated, under the assumption that they were part of the target population.

TABLE 5.1d

TEST LANGUAGE, TARGET POPULATION SIZE AND NUMBER OF RESPONDENTS (SECOND CYCLE)

Country	Test language	Population aged 16-65	Survey respondents aged 16-65
Chile	Spanish	9,400,000	3,502
Czech Republic	Czech	7,100,000	3,132
Denmark	Danish	3,400,000	3,026
Finland	Finnish	3,200,000	2,928
Hungary	Hungarian	7,000,000	2,593
Italy	Italian	38,700,000	2,974
Norway	Bokmål	2,800,000	3,307
Slovenia	Slovenian	1,400,000	2,972
Switzerland (Italian)	Italian	200,000	1,302

TABLE 5.1e

TOTAL NUMBER OF RESPONDENTS (FIRST AND SECOND CYCLES)

Survey cycle	Survey respondents aged 16–65[1]
First	36,533
Second	25,736
Total	**62,269**

1. Includes records where age was not stated, under the assumption that they were part of the target population.

5.2 Sample Designs

The IALS required all countries to employ a probability sample representative of the national population aged 16-65. No single sampling methodology was imposed due to differences in the data sources and resources available in each of the participating countries. A thorough review of the designs was conducted by Statistics Canada prior to the data collection operations to ensure that countries met the required sampling criteria. The second cycle required countries to supply more detailed sampling documentation. As such, a more comprehensive review was possible. The sample designs used by the participating countries are described below. Numbers of survey respondents refer to the full samples, see Tables 6.5a and 6.5b.

Sample Designs—First Cycle

Belgium (Flanders): The designated area of Flanders was divided into statistical sectors, from which 200 were selected with probability proportional to size. Then, 40 persons were chosen from a complete list of persons for each of these selected sectors. Finally, in order to get an equal distribution of persons by education level, the chosen persons were then selected into the final sample based on their level of education. Those people who were not sampled due to their education level were given a short questionnaire but these results were not included as part of the sample. This procedure explains, in part, the relatively low response rate achieved for the final sample. The total number of respondents was 2,261.

Canada: Two samples were combined. The main IALS sample was a sub-sample of the May 1994 Canadian Labour Force Survey (LFS) file using probability sampling at all stages. The sub-sample of 6,427 LFS respondents was stratified, with an over-sample of certain target groups of policy interest. The sample yielded 4,703 respondents. The other sample was a three-stage probability sample of Francophones from the province of Ontario selected from the 1991 census. This sample resulted in 1,044 respondents. The total number of respondents was 5,660.

Germany: The country used a master sample of sampling points, with the selection of addresses being made using the random route method. At each of the 525 sampling points, a single random route of addresses was followed, and along each route eight addresses were selected. In each household one person was selected for interview using the Kish method. The sample comprised 4,033 addresses, of which 997 did not belong to the target population. The total number of respondents was 2,062.

Ireland: Probability sampling was used at each of three stages of selection used. At the first stage of sampling, district electoral divisions were selected by stratum, where strata were defined in terms of population size and urban/rural type. Within each selected division, electoral registers were used to select a household. One adult per household was then selected randomly according to their date of birth. The total number of respondents was 2,423.

Netherlands: The Dutch approach was to use two-stage systematic sampling. In the first stage, postal codes were selected; in the second, one address was chosen from each selected postal code. The person to be interviewed in each sampled household was determined randomly according to their date of birth. The total number of respondents was 3,090.

New Zealand: The initial sampling frame was a list of geographical regions ("meshblocks"). The country was stratified by region and population size, and meshblocks were selected within strata with probability proportional to size. Households were then randomly selected within the meshblock. Finally, a Kish sampling grid was used to select one person per household. The total number of respondents was 4,223.

Poland: Poland used a stratified, multi-stage design employing probability sampling at the various stages. The sample was selected from the Polish National Register of Citizens, a register that covers all persons living permanently (longer than three months) in the country. The total number of respondents was 3,000.

Sweden: A stratified, self-weighting master sample was used. The sample was drawn from a national register of individuals. The total number of respondents was 3,038.

Switzerland (French and German): The target population was divided into two strata, corresponding to German- and French-speaking regions. Household telephone numbers were selected, and in each household the first member by alphabetical order of first name was selected. A complementary sample was selected in the canton of Geneva, using the same methods as the principal sample. The total number of respondents was 2,838.

United Kingdom: Two samples were selected – one for Great Britain and the other for Northern Ireland. In Great Britain, the Postal Code Address file was used to select the initial sample of addresses by postal code sectors. At each of the 35 addresses contained within each sector, the Kish method was used to select one adult. In Northern Ireland, a list of all private addresses was used to select an initial systematic sample of 7,000. At each of these addresses, one person was selected using the Kish method. The United Kingdom had a total sample of 6,718 respondents, 3,811 from Great Britain and 2,907 from Northern Ireland.

United States: The sample was selected from housing units undergoing their final Current Population Survey interviews during the period March-June 1994. A probability sample of 4,901 persons was selected using a disproportionate stratified design, with strata formed by race/ethnicity and education. This allocation was designed to provide an efficient linkage of the IALS survey to the earlier National Adult Literacy Survey (NALS). Students residing on college or university campus were excluded from the sample. The total number of respondents was 3,045.

All 12 first-cycle countries used probability sampling for most of the stages of their sample designs; in fact, ten used it in all stages. Two countries – Switzerland and Germany – used a non-probability sampling method in one stage of their multi-stage designs. Switzerland selected one household member using an alphabetic sort. This selection method is expected to yield unbiased results because of the unlikely correlation between first name and literacy skill level. Germany used the "random walk" method for selecting households for the sample. This non-probability method is often used with area frames because of practical constraints – namely the cost associated with enumerating every household within a geographic area, necessary for a probability sample. With non-probability sampling, there is no information about the properties of the resulting estimates, and so no definitive statement about their data quality can be made. This is not to say that the quality is better or worse than that of a probability sample; rather, the quality level is unknown. This issue is examined in greater detail in Murray *et al.* (1997).

Sample Designs – Second Cycle

Chile: A four-stage stratified sample design was used, with sampling units in a sequence extending from districts, census sectors, dwellings to individuals. Stratification of districts was performed according to region and type (urban/rural). Districts were selected with probability proportional to size in a systematic manner. In selected districts, census sectors were drawn again with probability proportional to size. A list of dwellings and individuals was drawn in those selected sectors during a preliminary visit. Dwellings were selected using the method of moving blocks and one individual in each selected dwelling was selected at random using a Kish table. Highly educated individuals were given a probability of selection twice as high as other individuals. The total number of respondents was 3,583.

Czech Republic: An area frame was used where primary sampling units were census units, defined as parts of cities, towns or villages with an average number of 80 households. Stratification of census units was done by size of locality and region. Selection of census units was carried out with probability proportional to the number of households, ensuring that there were at least two selections per stratum. The second stage of selection consisted in selecting an equal number of households in each unit, from an available list of households in those selected units. Finally, one individual was selected at random in each selected household, using a Kish table. The total number of respondents was 3,132.

Denmark: The sampling frame was the Population Register, which is kept up-to-date and includes all people living in Denmark. Individuals were directly selected from the frame at random. Stratification was done according to age and region. The total number of respondents was 3,026.

Finland: The sample of individuals was selected from the Central Population Register by systematic random sampling. The frame was sorted by a unique domicile code and by age. The sort order ensured implicit proportional stratification according to geographical population density. The total number of respondents was 2,928.

Hungary: The sampling frame was composed of two parts: a self-representing component (Budapest and the county seats) and the rest of the country. In the self-representing component, individuals were directly selected from the computerized database of the Central Office of Elections and Registration. In the rest of the country, stratification took place according to counties and size of settlements. The settlements themselves were the primary sampling units, selected with probability proportional to size. Individuals were then selected at random using the same database as for the self-representing component. The same number of individuals was selected in each settlement, which resulted in a self-weighting design. The total number of respondents was 2,593.

Italy: The sample design was a two-stage design in the larger cities and a three-stage design in the rest of the country. In the larger cities, a systematic sample of phone numbers was first selected from the phone directory, and one person at random was then selected in each household contacted. In the rest of the country, municipalities were stratified by region and used as primary sampling units, with selection taking place with probability proportional to size. Random selection of phone numbers and individuals was performed in those sampled municipalities, as in the larger cities. The total number of respondents was 2,974.

Norway: The sampling frame was composed of two parts: a self-representing component of municipalities with a population of 30,000 and over, and the rest of the country. In the self-representing component, individuals were directly selected from the Population Register in a systematic fashion. In the rest of the country, deep stratification took place first according to counties and second to a variety of characteristics. Primary sampling units were single municipalities or groups of municipalities, selected with probability proportional to size. At the second stage of selection (first in the self-representing component), individuals were further stratified according to their education, in order to over-sample individuals at both ends of the education spectrum. Individuals were selected using systematic sampling. The sample was supplemented by a special sample of job seekers, selected from a special register according to procedures similar to the main sample. The total number of respondents was 3,307.

Slovenia: The sample design was a two-stage cluster sample. The primary units were enumeration areas with an average size of 50 households. Stratification was performed implicitly as areas were sorted according first to regions and second to urban-rural type. Areas were selected with probability proportional to the number of eligible individuals. The Population Register maintained by the Ministry of Inner Affairs was used for the selection of individuals. Individuals in selected areas were sorted according to the street, house number, and family name and sampled in a systematic fashion. The total number of respondents was 2,972.

Switzerland (Italian): The sample covered only the Italian-speaking part of Switzerland, which includes the Canton of Ticino and the Italian-speaking regions of the Grison. A two-stage stratified sample of individuals was selected, where phone numbers represented the primary sampling units. Stratification of phone numbers took place according to statistical districts and territorial subdivisions. Individuals in contacted households were randomly selected according to pre-determined random numbers. The total number of respondents was 1,302.

The enhanced data quality procedures imposed during the second cycle of the survey led to the outcome that all countries implemented statistically sound sampling designs.

5.3

Overall Assessment of Data Quality

In-depth analysis of data quality issues was implemented in the second cycle for each country. Through these analyses a few problems were identified with certain methodologies that could potentially have an effect on international comparability. These problems are described below.

Hungary: Two problems occurred in Hungary. First, the response rate in Budapest was extremely low, at 26 per cent compared with 55 per cent or higher for other regions of the country. Second, the data suggest that quota sampling was used in rural areas. This gives rise to a concern about the probabilistic nature of the sample. As a consequence, the presence of bias with a non-negligible impact on the literacy estimates for the country cannot be ruled out.

Italy: In the case of Italy, they had such a low response rate (32.5%) that the possibility of bias in the results could not initially be eliminated within reasonable doubt. Italy fielded a large sample size (14,012) in order to obtain responses from some 3,000 individuals. Thus, even though post-stratification was performed in a state-of-the-art manner the possibility exists that respondents differ in their ability from non-respondents. It should be noted, however, that, upon analysis, the estimates and distribution of skills for Italy seem reasonable.

Norway (Bokmål): It appears that the replicate weights produced for the calculation of the precision of the estimates do not appropriately reflect the sample design used in the country. The complication is not related to the fact that the country fielded surveys of both national languages, one of Bokmål and the other of Nynorsk. As a consequence, variances, coefficients of variation and confidence intervals will be slightly underestimated. Estimates of literacy levels are not affected.

Switzerland (Italian): Switzerland has a rather low response rate (47 per cent). A non-response follow-up study indicated that the problem was due mainly to the selection of the sample of individuals. Analysis of the results revealed an over-representation of women, which called into question the random nature of the selection. An independent investigation carried out by the Swiss Federal Statistical Office confirmed that a sizeable proportion of interviewers had not properly followed the procedures for random selection. The Swiss Italian-speaking sample is considered to be somewhat biased in favour of people most likely to be at home during the day.

6.0
Data Collection and Processing

6.1
Introduction

The IALS gathered descriptive and proficiency information from sampled respondents through a background questionnaire and a series of assessment booklets containing prose, document, and quantitative literacy tasks. Survey respondents spent approximately 20 minutes answering a common set of background questions concerning their demographic characteristics, educational experiences, labor market experiences, and literacy related activities. Responses to these background questions make it possible to summarize the survey results using an array of descriptive variables, and also increase the accuracy of the proficiency estimates for various subpopulations. Background information was collected by trained interviewers.

After answering the background questions, the remainder of respondents' time was spent completing a booklet of literacy tasks designed to measure their prose, document, and quantitative skills. Most of these tasks were open-ended; that is, they required respondents to provide a written answer.

To achieve good content coverage of each of three literacy domains, the number of tasks in the assessment had to be quite large. Yet, the time burden for each respondent also needed to be kept within an acceptable range. To accommodate these two conflicting requirements—in other words, to reduce respondents' time burden without sacrificing good representation of the content domain—each respondent was administered only a fraction of the pool of tasks, using a variant of matrix sampling.

6.2
Data Collection and Processing

Data collection for the IALS project took place between 1994 and 1998, depending in which of the survey cycles a country participated. Tables 6.2a and 6.2b present the collection periods.

To ensure high quality data, the IALS Survey Administration Guidelines[7] specified that each country should work with a reputable data collection agency or firm, preferably one with its own professional, experienced interviewers. The manner in which these

7 For the IALS a large number of guidelines, technical specifications and other documents were written and made available to the national study teams in the participating countries. Examples are the IALS International Planning Report, the IALS Sampling Guidelines, the IALS Survey Administration Guidelines and the IALS Scoring Manual. These documents are available from the Special Surveys Division of Statistics Canada.

interviewers were paid should encourage maximum response. The interviews were conducted in homes in a neutral, non-pressuring manner. Interviewer training and supervision was to be provided, emphasizing the selection of one person per household (if applicable), the selection of one of the seven main task booklets (if applicable), the scoring of the core task booklet, and the assignment of status codes. Finally the interviewers' work was to have been supervised by using frequent quality checks at beginning of data collection, fewer quality checks throughout collection and having help available to interviewers during the data collection period.

TABLE 6.2a

SURVEY COLLECTION DATES (FIRST CYCLE)

Country	Collection date
Australia	May through July 1996
Belgium (Flanders)	1996
Canada	September through October 1994
Germany	September through November 1994
Ireland	1994
Netherlands	September through December 1994
New Zealand	1996
Poland	October 1994 through January 1995
Sweden	October 1994 through February 1995
Switzerland (French and German)	1994
United Kingdom	1996
United States	October through November 1994

TABLE 6.2b

SURVEY COLLECTION DATES (SECOND CYCLE)

Country	Collection date
Chile	May through June 1998
Czech Republic	December 1997 through March 1998
Denmark	April, May and August 1998
Finland	February until June 1998
Hungary	August through September 1998
Italy	September until December 1998
Norway	November 1997 through May 1998
Slovenia	September through November 1998
Switzerland (Italian)	March through September 1998

The IALS took several precautions against non-response bias, as specified in the IALS Administration Guidelines. Interviewers were specifically instructed to return several times to non-responding households in order to obtain as many responses as possible. In addition, all countries were asked to trace respondents who had moved, where applicable according to the sample design.

During the second cycle, data collection questionnaires were completed by study managers in order to demonstrate that the guidelines had been followed. Table 6.2c presents information about interviewers derived from these questionnaires.

TABLE 6.2c

INTERVIEWER INFORMATION (SECOND CYCLE)

Country	Number of interviewers	Number of supervisors	Interviewer experience
Chile	230	12	About one-half of the interviewers were university students in the social sciences and the rest were professional survey interviewers with an average experience of two years
Czech Republic	No information provided		
Denmark	112	5	Professional interviewers with between five and ten years of experience
Finland	135	3	Professional interviewers with on average 13 years of service
Hungary	150	5	Professional interviewers
Italy	180	8	Professional interviewers with 2 years experience and second level high school diploma
Norway	150	6	Professional interviewers with on average five years of survey experience
Slovenia	127	8	About 90 per cent were experienced in interviewing. The others had little or no survey experience
Switzerland (Italian)	56	2	Interviewers were trained especially for this survey

6.2.1

Model Procedures Manuals and Instruments

Each IALS country was given a set of administration manuals and survey instruments to use as a model. Countries were permitted to adapt these models to their own national data collection systems, but they were required to retain a number of key features. First, respondents were to complete the core and main test booklets alone, in their homes, without help from another person or from a calculator. Second, respondents were not to be given monetary incentives for participating. Third, despite the prohibition on monetary incentives, interviewers were provided with procedures to maximize the number of completed background questionnaires, and were to use a common set of coding specifications to deal with non-response. This last requirement is critical. Because non-completion of the core and main task booklets is correlated with ability, background information about non-respondents is needed in order to impute cognitive data for these persons.

6.2.2

Background questions

The model background questionnaires given to all IALS countries contained two sets of questions: mandatory questions, which all countries were required to include; and optional questions, which were recommended but not required. Countries were not required to field literal translations of the mandatory questions, but were asked to respect the conceptual intent of each question in adapting it for use. Countries were permitted to add questions to their background questionnaires if the additional burden on respondents would not reduce response rates. However, these questions will not be found in the international file, only in the country's national file.

6.2.3

Literacy tasks

The IALS is based on the premise that the difficulty of various literacy tasks is determined by certain factors, which are stable across language and culture. Accordingly, all of the IALS countries were given graphic files containing the pool of IALS literacy items and were instructed to modify each item by translating the English text to their own language without altering the graphic representation.

Certain rules governed the item modification process. For instance, some items required respondents to perform a task that was facilitated by the use of keywords. In some cases, the keywords were identical in the question and the body of the item; in others, the keyword was similar but not exactly the same; and in still other cases, the keyword was a synonym of the word used in the body of the item. In another case, respondents were asked to choose among multiple keywords in the body of the item, only one of which was correct. Countries were required to preserve these conceptual associations during the translation process.

Particular conventions used in the items—for example, currency units, date formats, and decimal delimiters—were adapted as appropriate for each country.

To ensure that the adaptation process did not compromise the psychometric integrity of the items, each country's test booklets were carefully reviewed for errors of adaptation.

6.2.4

Standardized non-response coding

It was crucial that the IALS countries managed non-respondent cases in a uniform manner so as to limit the level of non-response bias in the resulting survey estimates.

In IALS, a respondent had to complete the background questionnaire, pass the core block of literacy tasks, and attempt at least five tasks per literacy scale in order for researchers to be able to estimate his or her literacy skills directly. Literacy proficiency data were imputed for individuals who failed or refused to perform the core literacy tasks and for those who passed the core block but did not attempt at least five tasks per literacy scale. Because the model used to impute literacy estimates for non-respondents relies on a full set of responses to the background questions, IALS countries were instructed to obtain at least a background questionnaire from sampled individuals. They were also given a detailed non-response classification to use in the survey.

Each country was responsible for hiring its own interviewing staff. Thus, the number of interviewers, their pay rates, and the length of the survey period varied among the countries according to their norms and budgets. Each country was provided with a booklet to be used in training interviewers.

6.3

Scoring

Respondents' literacy proficiencies were estimated based on their performance on the cognitive tasks administered in the assessment. Unlike multiple-choice questions, which are commonly used in large-scale surveys and which offer a fixed number of answer choices, open-ended items such as those used in the IALS elicit a large variety of responses. Because raw data is seldom useful by itself, responses must be grouped in some way in order to summarize the performance results. As they were scored, responses to the IALS open-ended items were classified as correct, incorrect, or omitted.

The models employed to estimate ability and difficulty are predicated on the assumption that the scoring rubrics developed for the assessment were applied in a consistent fashion within and between countries. Several steps were taken to ensure that this assumption was met. Two of these main steps were the intra-country and inter-country rescores described in the following sections.

6.3.1

Intra-country rescoring

A variable sampling ratio procedure was set up to monitor scoring accuracy. At the beginning of scoring, almost all responses were rescored to identify inaccurate scorers and to detect unique or difficult responses that were not covered in the scoring manual. After a satisfactory level of accuracy was achieved, the rescoring ratio was dropped to a maintenance level to monitor the accuracy of all scorers. Average agreements were calculated across all items. To ensure that the first and second scores were truly independent, certain precautions had to be taken. For example, scorers had to be different persons, and the second scorer could not be able to see the scores given by the first scorer.

Scorers who received identical training within a country are expected to be more consistent amongst themselves than with scorers in other countries. Most of the rescoring reliabilities were above 97 per cent. It is important to note that the results were well within the statistical tolerances set for the IALS study and considerably better than those realized in other large-scale studies using open-ended items.

Since intra-country rescoring was used as a tool to improve data quality, score updates were not made to the database. In other words, the agreement data presented here indicate the minimum agreement achieved in scoring. After intra-country reliabilities were calculated, a few scorers were found to be unreliable. These scorers either received additional training or were released. Where scores and rescores differed, the first scores were replaced with correct scores if the inaccuracy was due to a systematic error on the part of the first scorer. In some cases, the scoring guide was found to be ambiguous. In such cases, the scoring guide was revised and the first scores were changed to reflect the revisions, but the second scores were not altered. The second scores were never replaced, even if they were subsequently found to be erroneous.

In sum, the first scores reflect changes and corrections resulting from lessons learned in the intra-country rescoring analysis. The first scores are therefore more accurate and consistent than the second scores, which retain errors and thereby underestimate the rescore reliabilities somewhat. The extent to which the reliabilities are underestimated must be very small, however, given that most of the reliabilities are above 97 per cent. These values indicate that very consistent scoring was achieved by all the participating countries.

6.3.2

Inter-country rescoring

Even after ensuring that all scorers were scoring consistently, fixing ambiguities in the scoring guides, and correcting any systematic scoring errors, it was still necessary to examine the comparability of scores across countries. Accurate and consistent scoring within a country does not necessarily imply that all countries are applying the scoring guides in the same manner. Scoring bias may be introduced if one country scores a certain response differently from the other countries. The inter-country rescorings described in this section were undertaken to ensure scoring comparability across countries.

As noted earlier, responses to the IALS assessment items were scored by each country separately. To determine inter-country scoring reliabilities for each item, the responses of a subset of examinees were scored by two separate groups. Usually, these scoring groups were from different countries. For example, a sample of test booklets was scored by two groups who scored Canada/English booklets and United States booklets. Inter-country score reliabilities were calculated by Statistics Canada, then evaluated by ETS. Based on the evaluation, every country was required to introduce a few minor changes in scoring procedures. In some cases, ambiguous instructions in the scoring manual were found to be causing erroneous interpretations and therefore lower reliabilities.

Using the inter-country score reliabilities, researchers can identify poorly constructed items, ambiguous scoring criteria, erroneous translations of items or scoring criteria, erroneous printing of items or scoring criteria, scorer inaccuracies, and, most important, situations in which one country consistently scored differently from another. In the latter circumstance, scorers in one country may consistently rate a certain response as being correct while those in another country score the same response as incorrect. This type of score asymmetry must be eliminated before the IRT scaling is performed. ETS and Statistics Canada identified such items, while the country in which the scoring problem occurred investigated the plausible causes for such systematic bias in scores. Where a systematic error was identified in a particular country, the original scores for that item were corrected for the entire sample.

Tables 6.3.2a and 6.3.2b summarize the inter-country rescore reliabilities, before corrections.

TABLE 6.3.2a

INTER-COUNTRY RESCORE RELIABILITY (FIRST CYCLE)

Original country	Number of booklets rescored	Average agreement (per cent)	Asymmetric items	Rescored by
Belgium (Flanders)	300	94		Netherlands
Canada/English	158	97	1	United States
Canada/French	142	97	7	France
Germany	270	94	6	Switzerland/German
Great Britain	300	97		Northern Ireland
Ireland	300	97		United States
Netherlands	300	96	2	Netherlands*
New Zealand	300	98		Australia
Northern Ireland	300	98		Great Britain
Poland	300	97	2	Canada
Sweden	300	97	1	Sweden*
Switzerland/French	154	96	11	France
Switzerland/German	153	96	4	Germany
United States	315	97	0	Canada/English

* The Netherlands and Sweden carried out both inter- and intra-country rescoring internally due to a lack of available language experts in Dutch and Swedish. Separate groups were established to perform the rescoring.

TABLE 6.3.2b

INTER-COUNTRY RESCORE RELIABILITY (SECOND CYCLE)

Original country	Number of booklets rescored	Average agreement (per cent)	Rescored by
Chile	349	92	Italy
Czech Republic	349	86	Canada
Denmark	350	88	Canada
Finland	354	95	Hungary
Hungary	346	91	Slovenia
Italy	350	93	Switzerland/Italian
Norway	346	90	Denmark
Slovenia	349	90	Canada
Switzerland (Italian)	327	80	Canada

Note: Inter-rater reliabilities for the IALS second cycle countries are high but generally somewhat lower than those for the IALS first cycle countries. This result is thought to be due to greater language heterogeneity in the second round of data collection.

6.4

Data Capture, Data Processing, and Coding

As a condition of their participation in the IALS, countries were required to capture and process their files using procedures that ensured logical consistency and acceptable levels of data capture error. Specifically, countries were advised to conduct complete verification of the captured scores (i.e., enter each record twice) in order to minimize error rates. Because the process of accurately capturing the test scores is essential to high data quality, 100 per cent keystroke validation was needed.

Each country was also responsible for coding industry, occupation, and education using standard international coding schemes (International Standard Industrial Classification, or ISIC 1968; International Standard of Classification Occupations, or ISCO 1988; and International Standard Classification of Education, or ISCED 1975). Further, coding schemes were provided for open-ended items, and countries were given specific instructions about the coding of such items so that coding error could be contained to acceptable levels.

In order to facilitate comparability in data analysis, each IALS country was required to map its national dataset into a highly structured, standardised record layout. In addition to specifying the position, format and length of each field, the international record layout included a description of each variable and indicated the categories and codes to be provided for that variable. Upon receiving a country's file, Statistics Canada performed a series of range checks to ensure compliance to the prescribed format. In the second cycle, Statistics Canada additionally ran consistency and flow edits on the data files received. When anomalies were detected, countries were notified of the problems and were asked to submit cleaned files.

6.5

Survey Response and Weighting

The IALS instrumentation consisted of three parts: (i) the background questionnaire, for demographic information about the respondent; (ii) the core tasks booklet, which identifies respondents with very low levels of literacy; and (iii) the main tasks booklet, used to calibrate the literacy levels.

The definition of an IALS respondent is a person who has fully or partially completed the background questionnaire. With this information, as well as the reason why the tasks booklet was not completed, it was possible to impute a literacy profile (given a sufficient number of complete responses). Thus the IALS procedures stressed that at a minimum the background questionnaire should be completed by every person sampled.
Tables 6.5a and 6.5b summarise the response rates achieved by the participating countries.

The variation in the number of out-of-scope cases in Table 6.5b appropriately reflects the fact that all countries used different sample designs.

At a meeting prior to the main data collection for the second cycle, countries had been asked to provide their overall sample size, the one that would yield the requested number of respondents (1,500 or 3,000) after non-response. They had also been advised against selecting additional samples in order to reach their target. The idea was to avoid any use of quota samples. Nonetheless, two countries – the Czech Republic and Denmark – did select additional samples. Given the small size of these additional samples (especially for Denmark), and the fact that satisfactory response rates had been obtained for both countries, it was felt that any potential impact of the additional samples would be negligible.

TABLE 6.5a

RESPONSE RATES BY COUNTRY (FIRST CYCLE)

Country	Age range	Number of respondents	Response rate (per cent)
Belgium (Flanders)[1]	16-65	2,261	36
Canada	16+	5,660	69
Germany	16-65	2,062	69
Ireland	16-65	2,423	60
Netherlands[1]	16-74	3,090	45
New Zealand	16-65	4,223	74
Poland[2]	16-65	3,000	75
Sweden	16+	3,038	60
Switzerland (French and German)	16+	2,838	55
United Kingdom	16-65	6,718	63
United States	16-65	3,045	60

1. Non-response follow-up surveys were conducted.
2. The response rate for Poland includes only the first wave of sampled persons, before interviewer follow-up.

TABLE 6.5b

RESPONSE RATES[1] BY COUNTRY (SECOND CYCLE)

Country	Age range	Initial sample	Additional sample	Total sample	Out-of-scope cases	Number of respondents	Response rate[1] (per cent)
Chile	15-65	5,200		5,200	384	3,583	74
Czech Republic	16-65	5,000	643	5,643	554	3,132	62
Denmark	16-65	4,500	115	4,615	9	3,026	66
Finland	16-65	4,250		4,250	10	2,928	69
Hungary	16-65	5,000		5,000	9	2,593	52
Italy	16-65			14,012[2]	4,872	2,974	33
Norway	16-65	5,500		5,500	68	3,307	61
Slovenia	16-65	4,290		4,290	12	2,972	69
Switzerland (Italian)	16-65	4,000		4,000	1,229	1,302	47

1. Calculated as the number of respondents divided by the total sample minus out-of-scope cases.
2. The exact breakdown of the Italian sample size is unknown.

The reason that low response rates are of concern in any survey is that bias might exist in the resulting estimates. Several precautions against non-response bias were taken. Interviewers were instructed to return several times to non-responding households in order to obtain as many responses as possible. In addition, all sample designs included some over-sampling. This refers to the inclusion in a sample of more randomly selected households than are necessary for the required number of completed interviews, to ensure a sufficient number of responses. Finally, the IALS sampling guidelines included an adjustment during the weighting procedure to help correct for non-response bias.

This correction, known as post-stratification, adjusts the population weights so that they match known population counts, *e.g.* by age group or education level. All countries post-stratified their data to such counts. The underlying assumption behind this compensation for non-response is that the respondents and non-respondents have the same literacy profile for the characteristic for which the adjustment is made. Tables 6.5c and 6.5d indicate the applied non-response adjustments.

TABLE 6.5c

POST-STRATIFICATION VARIABLES BY COUNTRY (FIRST CYCLE)

Country	Benchmark variables
Belgium	Age, sex, education
Canada	Province, economic region, census metropolitan area, age, sex, in-school youth, out-of-school youth, unemployment insurance recipients, social assistance recipients
Germany	Number of household members aged 16-65, age, sex, citizenship
Ireland	Area, sex, age
Netherlands	Region, age, sex, education
New Zealand	Sex, age, household size, urban/rural
Poland	Region, urban/rural, age
Sweden	Region, education, age, sex
Switzerland (French and German)	Number of household members aged 16-65, total number of persons in the household, level of education, size of community, age, sex
United Kingdom	Education, employment, region, age, sex
United States	Education

TABLE 6.5d

POST-STRATIFICATION VARIABLES BY COUNTRY (SECOND CYCLE)

Country	Benchmark variables (number of categories)
Chile	Urban/rural (2), age (5), sex (2)
Czech Republic	Education (4), age (3), sex (2), then region (8)
Denmark	Region (4), education (3), age (5), sex (2)
Finland	Province (6), education (5), age (5), sex (2), population density (3)
Hungary	Region (4), age (5), sex (2)
Italy	Region (4), urban/rural (2), education (5), age (5), sex (2)
Norway	Age (5), sex (2), then education (3)
Slovenia	Age (10), sex (2)
Switzerland (Italian)	Education (3), age (3), sex (2)

In the Czech Republic and Norway post-stratification was done in two stages. As a result, final estimates would not coincide perfectly with the benchmark totals obtained for the first group of variables.

7.0 Guidelines for Tabulation and Analysis

This section of the documentation outlines the guidelines to be adhered to by users tabulating, analysing, publishing or otherwise releasing any data derived from the survey microdata tapes. With the aid of these guidelines, users of microdata should be able to produce the same figures as those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

7.1 Sample Weighting Guidelines for Tabulation

The IALS surveys are based upon complex sample designs, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from bias, the survey weights must be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures differ from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are poor. Programs for calculating standard errors for simple estimates such as totals, proportions and ratios (for qualitative variables) are provided in the following section.

7.2 Definitions of Types of Estimates: Categorical vs. Quantitative

Before discussing how the IALS data can be tabulated and analyzed, it is useful to describe the two main types of point estimates of population characteristics, which can be generated from the microdata file for the IALS.

Categorical Estimates:
Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number of Germans at literacy Level 1 on the prose scale or the proportion of Canadians

at literacy Level 4 in numeracy are examples of such estimates. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of Categorical Questions:

Q: Do you ever watch television or videos in a language other than French or English?
R: Yes / No

Q: How would you rate your reading skills in English needed in daily life?
R: Excellent / Good / Moderate / Poor

Quantitative Estimates:
Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population. They also specifically involve estimates of the form $\hat{X}/\hat{Y}$ where $\hat{Y}$ is an estimate of surveyed population quantity total and $\hat{Y}$ is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is the average number of employers that working Canadians had in the past 12 months. The numerator is an estimate of the total number of employers that working Canadians had in the past 12 months, and its denominator is the number of Canadians reporting that they worked in the past 12 months.

Examples of Quantitative Questions:

Q: How many different employers have you had in the past 12 months?
R: |_|_| employer(s)

Q: How many hours per week did you usually work at this job?
R: |_|_| hours

7.2.1

Tabulation of Categorical Estimates

Estimates of the number of people within a given country with a certain characteristic can be obtained from the microdata file by summing the final weights of all records possessing the characteristic(s) of interest.

Proportions and ratios of the form $\hat{X}/\hat{Y}$ for a country are obtained by:

a) summing the final weights of records having the characteristic of interest for the numerator ($\hat{X}$),
b) summing the final weights of records having the characteristic of interest for the denominator ($\hat{Y}$), then
c) dividing the numerator estimate by the denominator estimate.

7.2.2

Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the microdata file by multiplying the value of the variable of interest by the final weight for each record, then summing this quantity over all records of interest. For example, to obtain an estimate for a particular country of the total number of different employers that people working part time have had in the past 12 months, multiply the value reported in the question D4 (number of employers) by the final weight for the record, then sum this value over all records with D5=2 (part time).

To obtain a weighted average of the form $\bar{X}/\bar{Y}$, the numerator ($\bar{X}$) is calculated as for a quantitative estimate and the denominator ($\bar{Y}$) is calculated as for a categorical estimate. For example, to estimate the average number of employers in the past 12 months of people working part time, in a given country

a) estimate the total number of employers as described above,
b) estimate the number of people in this category by summing the final weights of all records with QD5=2, then
c) divide estimate a) by estimate b).

7.3

Literacy Level Estimates

The IALS design is an adaptation of a three parameter logistic (PL) Item Response Theory model. The first parameter (A) is the ability of the item to discriminate (sensitivity to proficiency) and the second (B) is its difficulty. A third parameter (C) is the lower asymptote parameter which reflects the possibly non-zero chance of a correct response independent of ability. However, since the IALS test did not use any multiple choice type questions, this (C) parameter was fixed at zero throughout, thus transforming the equation into what can now be called a 2PL model. Once the parameters have been calculated, each item can be assigned a Response Probability value of 80 (RP80) which measures the proficiency level needed for a respondent to answer the task with an 80% probability of success. If four or more subpopulations displayed differential parameters, the item was dropped from the assessment and did not go into the calculation of the assessment of an individual's proficiency.

As noted previously, a respondent's proficiency in the three scales was summarized through the use of the item parameters and the respondent's ability in accordance with the IRT scaling models. The application differed from the norm in that the IALS called for administering relatively few items to each respondent in order to track population levels of proficiency more efficiently. Because the data are not intended to estimate individual levels of proficiency, however, more complicated analyses are required. Plausible values methodology was used to estimate key population features consistently and to approximate others no less accurately than standard IRT procedures would. In essence, this added dimension requires that the estimation of proficiency be based on a series of five plausible values for each of the three literacy domains. These five plausible values—prose1 through prose5 for the prose scale, doc1 through doc5 for the document scale

and quant1 through quant5 for the quantitative scale—have been recoded into plausible levels with values from 1 through 5 reflecting the empirically determined progression of information-processing skills and strategies required to perform increasingly complex tasks. Level 1 is equivalent to scores in the range 0 to 226 (inclusive); Level 2 is equal to scores of 226.0001 through 276; Level 3 goes from 276.0001 to 326; Level 4 includes scores ranging from 326.0001 to 376 and, Level 5 is equivalent to scores greater or equal to 376.0001. For the prose scale, the variables are called plev1 through plev5, for the document scale, these are dlev1 through dlev5 and for the quantitative scale, qlev1 through qlev5. Finally, in order to reproduce estimates published in the international and national Canadian (1996) reports, plev1 has been recoded into variable xprose whereby Levels 4 and 5 have been collapsed. Similarly, dlev1 has been recoded into xdoc and qlev1 into xquant. The reason for this recoding is to provide enough sample in each level to produce statistically meaningful estimates. The use of the first plausible value as the root for these estimates is entirely arbitrary and it would be equally legitimate to use any of the five values to produce point estimates. The table below demonstrates the inheritance tree for the plausible values, levels and reporting level for all three domains.

	PROSE1 →		PLEV1 →	1-4	XPROSE
	PROSE2 →		PLEV2		
0-500	PROSE3 →	1-5	PLEV3		
	PROSE4 →		PLEV4		
	PROSE5 →		PLEV5		
	DOC1 →		DLEV1 →	1-4	XDOC
	DOC2 →		DLEV2		
0-500	DOC3 →	1-5	DLEV3		
	DOC4 →		DLEV4		
	DOC5 →		DLEV5		
	QUANT1 →		QLEV1 →	1-4	XQUANT
	QUANT2 →		QLEV2		
0-500	QUANT3 →	1-5	QLEV3		
	QUANT4 →		QLEV4		
	QUANT5 →		QLEV5		

For simple point estimates in either of the three literacy domains, it is sufficient to use the population weight along with one of the corresponding five plausible values (chosen at random). To simplify this type of univariate or bivariate analysis, the variables xprose, xdoc and xquant are included on the international microdata file.

However, a more precise point estimate can be obtained by taking the average of the five estimates produced from each of the five plausible values, which can be computed as follows:

$$\boldsymbol{T.} = \left(\sum_i \boldsymbol{T_i}\right) / \mathbf{5}$$, where T_i is a vector of five weighted estimates derived from each of the five plausible values.

Note that taking an average of the five plausible values, will only produce a valid point estimate, not a valid variance estimate. **All five** plausible values as well as the 30 replicate weights must be used in order to correctly compute design-based variance estimates. Design-based variance estimates are discussed further in section 8.1.2. (Using Plausible Values and Replicate Weights in Calculating Sampling Errors).

7.4 Rounding Guidelines

In order that estimates for publication or other release derived from the microdata file correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.

b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.

c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e. numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding. In normal rounding to a single digit, if the final or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is increased by 1.

d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.

e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).

f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

8.0
Data Quality

The data quality from any survey can be evaluated by looking at two types of survey errors: sampling error and non-sampling error.

The estimates derived from this survey are based on a sample of individuals. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used. The difference between the estimates obtained from the sample and the results from a complete count taken under similar conditions is called the sampling error of the estimate.

Errors, which are not related to sampling, may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

8.1

Sampling Errors

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors is the standard error of the estimates derived from survey results.

However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (C.V.) of an estimate, is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose that, based upon the survey results, one estimates that 16.6% of Canadians are at literacy Level 1 with regard to prose, and this estimate is found to have standard error of 0.013. Then the coefficient of variation of the estimate is calculated as:

$$\left(\frac{.013}{.166}\right) x\ 100\% \ = \ 7.8\%$$

8.1.1

CV Release Guidelines

One criterion that can used to determine whether survey estimates are publishable is the coefficient of variation (CV). The CV is the standard error of an estimate expressed as a percentage of that estimate.

Before releasing and/or publishing any estimate from the IALS, users should first determine the quality level of the estimate. The quality levels are acceptable, marginal and unacceptable. Data quality is affected by both sampling and non-sampling errors. However for release purposes, the quality level of an estimate will be determined only on the basis of sampling error as reflected by the coefficient of variation as shown in table 8.1.1. Nonetheless users should be sure to read section 8 to be more fully aware of the quality characteristics of these data.

First, the number of respondents who contribute to the calculation of the estimate should be determined. If this number is less than 30, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 30 or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to weighted rounded estimates.

All estimates can be considered releasable. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

Table 8.1.1: Quality Level Guidelines

Quality level of estimate	Guidelines
1. Acceptable	Estimates have: a sample size of 30 or more, and low coefficients of variation in the range 0.0% to 16.5%. No warning is required.
2. Marginal	Estimates have: a sample size of 30 or more, and high coefficients of variation in the range 16.6% to 33.3%. Estimates should be flagged with the letter M (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error associated with the estimates.
3. Unacceptable	Estimates have: a sample size of less than 30, or very high coefficients of variation in excess of 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates: “The user is advised that . . . (specify the data) . . . do not meet Statistics Canada’s quality standards for this statistical program. Conclusions based on these data will be unreliable, and most likely invalid. These data and any consequent findings should not be published. If the user chooses to publish these data or findings, then this disclaimer must be published with the data.”

8.1.2

Using Plausible Values and Replicate Weights in Calculating Sampling Error

IALS countries used a variety of sampling schemes depending upon what was most efficient in each country. Thus, the jackknife technique has been chosen as an appropriate variance estimation technique due to its ability to handle various complex sampling designs. Using a jackknife variance estimator allows for fairly precise estimates of the total sampling error for population estimates and for conducting multivariate analyses. The jackknife procedure has a number of properties that make it particularly suited to the analysis of these data:

a) It provides unbiased estimates of the sampling error arising from the complex sample selection procedure for linear estimates such as simple totals and means, and does so approximately for more complex estimates.

b) It reflects the component of sampling error introduced by the use of weighting factors, such as non-response adjustments, that are dependent on the sample data actually obtained.

c) It can be adapted readily to the estimation of sampling errors for parameters estimated using statistical modelling procedures, as well as for tabulation estimates such as totals and means.

d) Once appropriate weights are derived and attached to each record, jackknifing can be used to estimate sampling errors. A single set of replicate weights is required for all tabulations and model parameter estimates that may be needed.

When computing jackknife variance estimates for literacy score estimates, it is important to use all five plausible values in the equation as well as the thirty replicate weights. This is a cumbersome procedure requiring the replication of tabulations using each of the replicate weights and each of the plausible values. In effect, for each variance estimate required, five sets of thirty pseudo estimates (5 plausible values X 30 replicate weights) must be produced. The first component of the variance formula is the mean of the five variances computed from each of the five sets of thirty pseudo estimates. The second component of the variance formula, which is multiplied by a factor of 6/5 and added to the first component, is the variance of the five estimates produced from each of the five plausible values. The formula is as follows[8]:

$$\mathbf{Var}(T_{\cdot}) = \left[\sum_i \left(\left(\sum_j (\dagger_{ij} - T_i)^2\right)/(30*29)\right)\right]/5 + (6/5)\sum_i (T_i - T_{\cdot})^2/4$$

Where,
i = 1,…,5 represents the five plausible values,
j = 1,…,30 represents the thirty replicate weights,

8. The jackknife variance formula combines the variance arising from sampling respondents and the variance arising from the modelling of proficiency scores. The first component of the right-hand side approximates the sampling error due to sampling design (Wolter, 1985, p. 183); the second component approximates the error due to the model (Murray, Kirsch, Jenkins, 1998, p.182).

$\mathbf{T.} = (\sum_i \mathbf{T}_i) / 5$

Where $\mathbf{T}_i$ represents each of the five estimates derived from each of the five plausible values using the final weight.

$\mathbf{T}_i = \sum_j \mathbf{t}_{ij} / 30$

Where $\mathbf{t}_{ij}$ represents each of 150 pseudo values derived as follows:

$\mathbf{t}_{ij} = (30 * \mathbf{T}_i) - (29 * \mathbf{t}_{ij})$,

Where $\mathbf{t}_{ij}$ represents each of 150 estimates derived from each of the five plausible values times thirty replicate weights.

The correct standard error is the square root of Var (T.). Such standard errors would include errors from both sampling and modelling.

Calculating Variance Estimates for Simple Means and Totals

It is possible to do the jackknifing procedure using SPSS or SAS in a single pass. The following routines detail the procedure for calculating variance estimates for simple means and totals. The routines for linear estimates such as simple totals and means provide unbiased estimates of the sampling error arising from the complex sample selection procedures.

Programs I-IV are written in SPSS and programs V-VIII are written in SAS. Examples of what the output should look like are provided after each program.

For the following programs note:
The bold characters are the only variables and strings that need to be modified.
It is crucial that the case weights not be applied to the datasets before executing any of the following procedures, since the procedures themselves weight the data. If this is done, the estimates produced will be incorrect because the weight will have been applied twice. In addition, these procedures do not handle missing values, as do other predefined procedures within SPSS or SAS. Thus, whenever it is necessary, make sure to exclude cases with missing values that may affect the final results.

I. Standard error and mean computation for literacy scores: Multiweight method using SPSS with correction for imputation

This program provides mean literacy scores and the associated standard errors for any of the three literacy scales. The estimates can be produced for any categorical break variable or a combination of categorical break variables. The following example produces mean scores and standard errors on the prose scale for each gender within each country.

```
Get File= 'path and filename of dataset'
        /Keep= cntrid (or other break variable), gender (or other break variable(s)), prose1
        To prose5 (or doc1 To doc5 or quant1 To quant5), age, weight, replic01 To replic30.
Weight off.
Select if ((age>=16 and age<=65) or age=9).
Select if (not sysmis(cntrid)).
Select if (not sysmis(gender)).

Vector WT= replic01 To replic30.
Vector AWX(30).
Vector BWX(30).
Vector CWX(30).
Vector DWX(30).
Vector EWX(30).
Loop #i= 1 To 30.
        Compute AWX(#i)= WT(#i)*prose1.
        Compute BWX(#i)= WT(#i)*prose2.
        Compute CWX(#i)= WT(#i)*prose3.
        Compute DWX(#i)= WT(#i)*prose4.
        Compute EWX(#i)= WT(#i)*prose5.
End Loop.

Vector VALUE= prose1 To prose5.
Vector WS(5).
Loop #i= 1 To 5.
        Compute WS(#i)= VALUE(#i)*weight.
End Loop.
Execute.

Aggregate Outfile= 'path and filename of aggregate file to create' /Break=cntrid gender/
UNW=N(weight)/SWT, SW1 To SW30= Sum(weight, replic01 To replic30)/
ASX1 To ASX30, BSX1 To BSX30, CSX1 To CSX30, DSX1 To DSX30, ESX1 To ESX30
=Sum(AWX1 To AWX30, BWX1 To BWX30, CWX1 To CWX30, DWX1 To DWX30,
EWX1 To EWX30)/SS1 TO SS5= Sum(WS1 To WS5)/.

Get File= 'path and filename of aggregate file created'.

Vector SA= SW1 To SW30.
Vector SB= SW1 To SW30.
Vector SC= SW1 To SW30.
Vector SD= SW1 To SW30.
Vector SE= SW1 To SW30.
Vector VSX= ASX1 To ASX30.
Vector WSX= BSX1 To BSX30.
Vector XSX= CSX1 To CSX30.
Vector YSX= DSX1 To DSX30.
Vector ZSX= ESX1 To ESX30.
Vector AXBAR(30).
Vector BXBAR(30).
Vector CXBAR(30).
Vector DXBAR(30).
Vector EXBAR(30).
```

{The bold characters are the only variables and strings that need modification}

```
Loop #i= 1 To 30.
        Compute AXBAR(#i)= VSX(#i)/SA(#i).
        Compute BXBAR(#i) = WSX(#i)/SB(#i).
        Compute CXBAR(#i) = XSX(#i)/SC(#i).
        Compute DXBAR(#i) = YSX(#i)/SD(#i).
        Compute EXBAR(#i) = ZSX(#i)/SE(#i).
End Loop.

Vector SI= SS1 To SS5.
Vector STI(5).
Loop #i= 1 To 5.
        Compute STI(#i)= SI(#i)/SWT.
End Loop.

Compute XBAR= Mean(STI1 To STI5).

Vector PAXBAR(30).
Vector PBXBAR(30).
Vector PCXBAR(30).
Vector PDXBAR(30).
Vector PEXBAR(30).
Loop #i= 1 To 30.
        Compute PAXBAR(#i)=(30*XBAR) - (29*AXBAR(#i)).
        Compute PBXBAR(#i)=(30*XBAR) - (29*BXBAR(#i)).
        Compute PCXBAR(#i)=(30*XBAR) - (29*CXBAR(#i)).
        Compute PDXBAR(#i)=(30*XBAR) - (29*DXBAR(#i)).
        Compute PEXBAR(#i)=(30*XBAR) - (29*EXBAR(#i)).
End Loop.

Compute SVAR1= Variance(PAXBAR1 To PAXBAR30)/30.
Compute SVAR2= Variance(PBXBAR1 To PBXBAR30)/30.
Compute SVAR3= Variance(PCXBAR1 To PCXBAR30)/30.
Compute SVAR4= Variance(PDXBAR1 To PDXBAR30)/30.
Compute SVAR5= Variance(PEXBAR1 To PEXBAR30)/30.

Compute SVAR= Mean(SVAR1 To SVAR5).
Compute MVAR= Variance(STI1 To STI5).
Compute SE= SQRT(SVAR+(6/5)*MVAR).
Compute CV = (SE/XBAR)*100.
Execute.

Print Formats SVAR, MVAR, XBAR, SE, CV (F8.4).
List cntrid gender UNW SWT XBAR SE CV.
OR {if using TABLES option}
Tables
      /Observation = UNW SWT XBAR SE CV
      /Table =  cntrid > gender BY (UNW + SWT + XBAR + SE + CV)
      /Title 'Country and Gender by Mean Prose Scores'.
```

The final output will have two lines for each country, one for males and the other for females. The variable XBAR provides the mean score and the variable SE provides the standard error for the mean estimate. CV provides the coefficient of variation. The variable SWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SPSS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program I:

Country and Gender by Mean Prose Scores						
		UNW	SWT	XBAR	SE	CV
Germany	Men	938	26874222	276.8	2.0	0.7
	Women	1124	26952066	274.9	1.4	0.5
United States	Men	1434	75312132	269.3	2.3	0.8
	Women	1601	84983108	277.6	2.2	0.8
Ireland	Men	1077	1092200	262.9	5.3	2.0
	Women	1346	1082180	268.4	2.3	0.9
Netherlands	Men	1358	5325766	281.5	1.3	0.5
	Women	1479	5134592	283.9	1.4	0.5
Poland	Men	1431	12130543	227.9	1.1	0.5
	Women	1569	12345106	231.0	1.7	0.7
Sweden	Men	1289	2667574	300.6	1.8	0.6
	Women	1355	2692237	301.9	1.7	0.6

II. Standard error and ratio computation for literacy levels: Multiweight method using SPSS with correction for imputation

This program computes proportions at each literacy level and the associated standard errors for any of the three literacy scales. The estimates can be produced for any categorical break variable or a combination of categorical break variables. The following example produces the proportion of individuals at each prose literacy level along with their associated standard errors within each country.

Get File= **'path and filename of dataset'**
/Keep= **cntrid (or other break variable(s)), plev1 To plev5 (or dlev1 To dlev5 or qlev1 To qlev5)**, age, weight, replic01 To replic30.
Weight off.
Select if ((age>=16 and age<=65) or age=9).
Select if (not sysmis(**cntrid**)).
Recode **plev1 plev2 plev3 plev4 plev5** (5=4).

Vector WT= replic01 To replic30.
Aggregate outfile= **'path and filename for first aggregate file created'**/Break=**cntrid**/
unw=N(Weight)/SWT, SW1 To SW30= Sum(Weight, replic01 To replic30)/.
Aggregate outfile= **'path and filename for second aggregate file created'**/Break=**cntrid plev1**/unw1=N(Weight)/VSWT, VSW1 To VSW30= Sum(Weight, replic01 To replic30).
Aggregate outfile= **'path and filename for third aggregate file created'**/Break=**cntrid plev2**/unw2=N(Weight)/WSWT, WSW1 To WSW30= Sum(Weight, replic01 To replic30)/.
Aggregate outfile= **'path and filename for fourth aggregate file created'** /Break=**cntrid plev3**/unw3=N(Weight)/XSWT, XSW1 To XSW30= Sum(Weight, replic01 To replic30)/.
Aggregate outfile= **'path and filename for fifth aggregate file created'**/Break=**cntrid plev4**/unw4=N(Weight)/YSWT, YSW1 To YSW30= Sum(Weight, replic01 To replic30)/.
Aggregate outfile= **'path and filename for sixth aggregate file created'** /Break=**cntrid plev5**/unw5=N(Weight)/ZSWT, ZSW1 To ZSW30= Sum(Weight, replic01 To replic30)/.

```
Get file='path and filename for second aggregate file created'.
Save outfile='path and filename for second aggregate file created'/rename plev1=plev.
Get file='path and filename for third aggregate file created'.
Save outfile='path and filename for third aggregate file created'/rename plev2=plev.
Get file='path and filename for fourth aggregate file created'.
Save outfile='path and filename for fourth aggregate file created'/rename plev3=plev.
Get file='path and filename for fifth aggregate file created'.
Save outfile='path and filename for fifth aggregate file created'/rename plev4=plev.
Get file='path and filename for sixth aggregate file created'.
Save outfile='path and filename for sixth aggregate file created'/rename plev5=plev.

MATCH FILES /FILE= 'path and filename for second aggregate file created'
 /TABLE= 'path and filename for first aggregate file created'
 /BY cntrid.
Save Outfile= 'path and filename for first merged file to create'.
MATCH FILES /FILE= 'path and filename for third aggregate file created'
 /TABLE= 'path and filename for first merged file created'
 /BY cntrid plev.
Save Outfile= 'path and filename for second merged file to create'.
MATCH FILES /FILE= 'path and filename for fourth aggregate file created'
 /TABLE= 'path and filename for second merged file created'
 /BY cntrid plev.
Save Outfile= 'path and filename for third merged file to create'.
MATCH FILES /FILE= 'path and filename for fifth aggregate file created'
 /TABLE= 'path and filename for third merged file created'
 /BY cntrid plev.
Save Outfile= 'path and filename for fourth merged file to create'.
MATCH FILES /FILE= 'path and filename for sixth aggregate file created'
/TABLE= 'path and filename for fourth merged file created'
/BY cntrid plev.
Save Outfile= 'path and filename for fifth merged file to create'
        /Keep=cntrid, plev, unw, unw1, unw2, unw3, unw4, unw5, SWT, VSWT, WSWT, XSWT,
        YSWT, ZSWT, SW1 to SW30, VSW1 to VSW30, WSW1 to WSW30, XSW1 to XSW30,
        YSW1 to YSW30, ZSW1 to ZSW30.

Get file= 'path and filename for fifth merged file created'.

Compute XBAR1= (VSWT/SWT).
Compute XBAR2= (WSWT/SWT).
Compute XBAR3= (XSWT/SWT).
Compute XBAR4= (YSWT/SWT).
Compute XBAR5= (ZSWT/SWT).
Compute XBAR= MEAN(XBAR1 to XBAR5).

Vector VSW= VSW1 To VSW30.
Vector WSW= WSW1 To WSW30.
Vector XSW= XSW1 To XSW30.
Vector YSW= YSW1 To YSW30.
Vector ZSW= ZSW1 To ZSW30.
Vector SW= SW1 To SW30.
Vector AXBAR(30).
Vector BXBAR(30).
Vector CXBAR(30).
Vector DXBAR(30).
Vector EXBAR(30).
```

{The bold characters are the only variables and strings that need modification}

```
Loop #i= 1 To 30.
      Compute AXBAR(#i)= ((VSW(#i)/SW(#i))).
      Compute BXBAR(#i)= ((WSW(#i)/SW(#i))).
      Compute CXBAR(#i)= ((XSW(#i)/SW(#i))).
      Compute DXBAR(#i)= ((YSW(#i)/SW(#i))).
      Compute EXBAR(#i)= ((ZSW(#i)/SW(#i))).
End loop.

Vector PAXBAR(30).
Vector PBXBAR(30).
Vector PCXBAR(30).
Vector PDXBAR(30).
Vector PEXBAR(30).
Vector AXBAR = AXBAR1 to AXBAR30.
Vector BXBAR = BXBAR1 to BXBAR30.
Vector CXBAR = CXBAR1 to CXBAR30.
Vector DXBAR = DXBAR1 to DXBAR30.
Vector EXBAR = EXBAR1 to EXBAR30.
Loop #i= 1 To 30.
         Compute PAXBAR(#i)=(30*XBAR) - (29*AXBAR(#i)).
         Compute PBXBAR(#i)=(30*XBAR) - (29*BXBAR(#i)).
         Compute PCXBAR(#i)=(30*XBAR) - (29*CXBAR(#i)).
         Compute PDXBAR(#i)=(30*XBAR) - (29*DXBAR(#i)).
         Compute PEXBAR(#i)=(30*XBAR) - (29*EXBAR(#i)).
End Loop.

Compute SVAR1= Variance(PAXBAR1 To PAXBAR30)/30.
Compute SVAR2= Variance(PBXBAR1 To PBXBAR30)/30.
Compute SVAR3= Variance(PCXBAR1 To PCXBAR30)/30.
Compute SVAR4= Variance(PDXBAR1 To PDXBAR30)/30.
Compute SVAR5= Variance(PEXBAR1 To PEXBAR30)/30.

Compute SVAR= Mean(SVAR1 To SVAR5).
Compute MVAR= Variance(XBAR1 To XBAR5).
Compute SE= SQRT(SVAR+(6/5)*MVAR).

Compute SVARpct=SVAR*100.
Compute MVARpct=MVAR*100.
Compute XBARpct=XBAR*100.
Compute SEpct= SE*100.
Compute CVpct= (SEpct/XBARpct)*100.
Execute.

Print Formats SVARpct, MVARpct, XBARpct, SEpct, CVpct (F8.4).
List cntrid plev UNW unw1 unw2 unw3 unw4 unw5.
List cntrid plev SWT vswt wswt xswt yswt zswt.
List cntrid plev XBARpct SEpct CVpct.
```

The final output will have four lines for each country, one for each prose level. The variable XBARpct provides the proportion of individuals at a given level within a country and the variable SEpct provides the standard error for the proportion. CVpct provides the coefficient of variation. The variable VSWT gives a weighted cell count for the first plausible value and UNW1 provides an unweighted cell count for the first plausible value. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SPSS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program II:

Country by Prose Level						
	Prose Scale	UNW1	VSWT	XBARpct	SEpct	CVpct
Germany	Level 1	299	7724776	13.8	1.3	9.5
	Level 2	715	18424267	35.3	1.4	3.9
	Level 3	768	20439707	37.3	1.7	4.5
	Level 4/5	280	7237538	13.6	1.0	7.1
United States	Level 1	843	33117817	20.8	1.2	5.7
	Level 2	778	41438190	24.5	1.5	6.0
	Level 3	881	51895358	32.8	1.3	3.9
	Level 4/5	533	33843876	21.9	1.3	5.9
Ireland	Level 1	557	491423.3	22.5	1.6	7.3
	Level 2	738	647543.1	30.5	1.6	5.3
	Level 3	812	742544	33.8	1.7	5.0
	Level 4/5	316	292869.7	13.2	1.7	12.6
Netherlands	Level 1	256	1101215	10.4	0.7	6.5
	Level 2	812	3147855	29.4	1.0	3.4
	Level 3	1292	4612593	44.7	1.4	3.1
	Level 4/5	477	1598696	15.5	1.1	7.1
Poland	Level 1	1288	10438093	42.7	0.8	1.9
	Level 2	1036	8455395	34.3	0.9	2.6
	Level 3	587	4834112	19.2	0.8	4.1
	Level 4/5	89	748048.4	3.7	0.6	17.1
Sweden	Level 1	192	401835.1	7.2	0.6	8.0
	Level 2	530	1089885	20.7	0.8	3.8
	Level 3	1059	2129211	39.8	1.0	2.5
	Level 4/5	863	1738880	32.2	1.0	3.2

III. Standard error computation for quantitative variables excluding literacy scores: Multiweight method using SPSS.

This program computes standard errors for quantitative variables other than the plausible values (i.e. other than the literacy scores). The mean for variables such as duration of training (Derived by multiplying variables F8M1*F9M1*F10M1, gives duration of first mentioned course/program) or other continuous variables that may be derived can have their means calculated along with the standard error in the following program. The following example produces mean estimates for years of education and their associated standard errors for each gender within each country.

Get File= **'path and filename of dataset'**
/Keep= **cntrid, gender (or other break variable(s))**, **a quantitative variable (e.g. years of education – a7),** age, weight, replic01 To replic30.
Weight off.
Select if ((age>=16 and age<=65) or age=9).
Select if (not sysmis(**gender**)).
Recode a7 (98,99=sysmis) (else=copy).
Select if (not sysmis(**a7**)).

```
Vector WT= replic01 To replic30.                    {The bold characters are
Vector WX(30).                                      the only variables and strings that
Compute WTX= Weight*a7.                             need modification}
Loop #i= 1 To 30.
   Compute WX(#i)= WT(#i)*a7.
End loop.
Aggregate outfile= 'path and filename of aggregate file to create' /Break=cntrid
gender/UNW=N(weight)/SWT, SW1 To SW30= Sum(weight, replic01 To replic30)/SWX, SX1 To
SX30= Sum(WTX, WX1 To WX30)/.

Get File= 'path and filename of aggregate file created'.

Vector SA= SW1 To SW30.
Vector SX= SX1 To SX30.
Vector AXBAR(30).
Loop #i= 1 To 30.
        Compute AXBAR(#i)= ASX(#i)/SA(#i).
End Loop.

Compute XBAR= SWX/SWT.

Vector PAXBAR(30).
Loop #i= 1 To 30.
        Compute PAXBAR(#i)=(30*XBAR) - (29*AXBAR(#i)).
End Loop.

Compute SVAR= Variance(PAXBAR1 To PAXBAR30)/30.
Compute SE= SQRT(SVAR).
Compute CV = (SE/XBAR)*100.
Execute.

Print Formats SVAR, XBAR, SE, CV (F8.4).
List cntrid gender UNW SWT XBAR SE CV.
```

The final output will have two lines for each country, one for males and the other for females. The variable XBAR provides the mean years of education and the variable SE provides the standard error for the mean estimate. CV provides the coefficient of variation. The variable SWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SPSS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program III:

Country and Gender by Years of Education						
		UNW	SWT	XBAR	SE	CV
Germany	Men	931	26734502	11.6	0.1	1.2
	Women	1110	26559411	11.1	0.1	0.8
United States	Men	1355	70826370	13.4	0.1	0.6
	Women	1530	81363215	13.0	0.1	0.6
Ireland	Men	1075	1090292	10.3	0.2	1.6
	Women	1341	1078606	10.5	0.1	0.9
Netherlands	Men	1358	5325766	13.1	0.1	0.8
	Women	1479	5134592	12.2	0.1	0.8
Poland	Men	1430	12120961	11.0	0.1	0.7
	Women	1558	12259003	10.9	0.1	0.6
Sweden	Men	1286	2661256	11.7	0.1	0.7
	Women	1352	2686417	11.6	0.1	0.8

IV. Standard error computation for background (categorical) variables: Multiweight method using SPSS:

This program computes proportions and the associated standard errors using the thirty replicate weights for any two (or more, if additional break variables are added) categorical background variables. The following example produces the proportions of males and females along with their associated standard errors within each country.

```
Get File= 'path and filename of dataset'
        /Keep= cntrid (or other break variable), gender (or other break variable(s)), age,
        weight, replic01 To replic30.
Weight off.
Select if ((age>=16 and age<=65) or age=9).
Select if (not sysmis(cntrid)).                    {The bold characters are
Select if (not sysmis(gender)).                    the only variables and strings that
                                                   need modification}
Vector WT= replic01 To replic30.
Aggregate outfile= 'path and filename for first aggregate file to create' /Break=cntrid
/UNWT=N(weight)/SWT, SW1 To SW30= Sum(weight, replic01 To replic30)/.
Aggregate outfile= 'path and filename for second aggregate file to create' /Break=cntrid
gender/UNW=N(weight)/ZSWT, ZSW1 To ZSW30= Sum(weight, replic01 To replic30)/.

Match Files /File= 'path and filename for first aggregate file created'
 /Table= 'path and filename for second aggregate file created'
 /By cntrid.
Save Outfile= 'path and filename for merged aggregate file created'.

Get file= 'path and filename for merged aggregate file created'.

Compute XBAR= (ZSWT/SWT).
```

```
Vector ZSW= ZSW1 To ZSW30.
Vector SW= SW1 To SW30.
Vector AXBAR(30).
Loop #i= 1 To 30.
      Compute AXBAR(#i)= ((ZSW(#i)/SW(#i))).
End loop.

Vector PAXBAR(30).
Loop #i= 1 To 30.
         Compute PAXBAR(#i)=(30*XBAR) - (29*AXBAR(#i)).
End Loop.

Compute SVAR= Variance(PAXBAR1 To PAXBAR30)/30.
Compute SE= SQRT(SVAR).
Compute SVARpct=SVAR*100.
Compute XBARpct=XBAR*100.
Compute SEpct= SE*100.
Compute CVpct= (SEpct/XBARpct)*100.
Execute.

Print Formats SVARpct, XBARpct, SEpct, CVpct (F8.4).
List cntrid gender UNW SWT XBARpct SEpct CVpct.
```

The final output will have two lines for each country, one for males and the other for females. The variable XBARpct provides the proportion of individuals within each gender type and the variable SEpct provides the standard error for the proportion. CVpct provides the coefficient of variation. The variable ZSWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SPSS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program IV:

Country by Gender						
		UNW	ZSWT	XBARpct	SEpct	CVpct
Germany	Men	938	26874222	49.9	1.2	2.4
	Women	1124	26952066	50.1	1.2	2.4
United States	Men	1434	75312132	47.0	1.1	2.3
	Women	1601	84983108	53.0	1.1	2.0
Ireland	Men	1077	1092200	50.2	1.0	2.0
	Women	1346	1082180	49.8	1.0	2.0
Netherlands	Men	1358	5325766	50.9	0.1	0.2
	Women	1479	5134592	49.1	0.1	0.2
Poland	Men	1431	12130543	49.6	0.0	0.0
	Women	1569	12345106	50.4	0.0	0.0
Sweden	Men	1289	2667574	49.8	1.2	2.4
	Women	1355	2692237	50.2	1.2	2.4

The next four programs are identical to the preceding four, but are written in SAS language.

V. Standard error and mean computation for literacy scores: Multiweight method using SAS with correction for imputation

This program provides mean literacy scores and the associated standard errors for any of the three literacy scales. The estimates can be produced for any categorical break variable or a combination of categorical break variables. The following example produces mean scores and standard errors on the prose scale for each gender within each country.

```
Data A;
Set libname.filename (keep= cntrid (or other break variable) gender (or other break
variable(s)) prose1--prose5 (or doc1 To doc5 or quant1 To quant5) age weight replic01--
replic30);
if ((age >=16 and age <=65) or age=9);

Array WT replic01-replic30;
Array AWX AWX1-AWX30;
Array BWX BWX1-BWX30;
Array CWX CWX1-CWX30;
Array DWX DWX1-DWX30;                    {The bold characters are
Array EWX EWX1-EWX30;                    the only variables and strings that
Do Over WT;                              need modification}
AWX = WT*prose1;
BWX = WT*prose2;
CWX = WT*prose3;
DWX = WT*prose4;
EWX = WT*prose5;
end;

Array VALUE prose1-prose5;
Array WS WS1-WS5;
Do Over WS;
WS = VALUE*weight;
end;

Proc Summary Data=A;
Class cntrid gender;
Var weight replic01-replic30 AWX1-AWX30 BWX1-BWX30
CWX1-CWX30 DWX1-DWX30 EWX1-EWX30 WS1-WS5;
Output Out=B N(weight)=UNW
Sum(weight replic01-replic30 AWX1-AWX30 BWX1-BWX30
CWX1-CWX30 DWX1-DWX30 EWX1-EWX30 WS1-WS5)=
SWT SW1-SW30 ASX1-ASX30 BSX1-BSX30
CSX1-CSX30 DSX1-DSX30 ESX1-ESX30 SS1-SS5;
```

```
Data C;
Set B;
Array SW SW1-SW30;
Array VSX ASX1-ASX30;
Array WSX BSX1-BSX30;
Array XSX CSX1-CSX30;
Array YSX DSX1-DSX30;
Array ZSX ESX1-ESX30;
Array AXBAR AXBAR1-AXBAR30;
Array BXBAR BXBAR1-BXBAR30;
Array CXBAR CXBAR1-CXBAR30;
Array DXBAR DXBAR1-DXBAR30;
Array EXBAR EXBAR1-EXBAR30;
Do over SW;
AXBAR = VSX/SW;
BXBAR = WSX/SW;
CXBAR = XSX/SW;
DXBAR = YSX/SW;
EXBAR = ZSX/SW;
end;

Array SS SS1-SS5;
Do Over SS;
SS = SS/SWT;
end;

XBAR = Mean(Of SS1-SS5);

Array PAXBAR PAXBAR1-PAXBAR30;
Array PBXBAR PBXBAR1-PBXBAR30;
Array PCXBAR PCXBAR1-PCXBAR30;
Array PDXBAR PDXBAR1-PDXBAR30;
Array PEXBAR PEXBAR1-PEXBAR30;
Do over AXBAR;
PAXBAR = (30*XBAR) - (29*AXBAR);
PBXBAR = (30*XBAR) - (29*BXBAR);
PCXBAR = (30*XBAR) - (29*CXBAR);
PDXBAR = (30*XBAR) - (29*DXBAR);
PEXBAR = (30*XBAR) - (29*EXBAR);
end;

SVAR1 = Var(Of PAXBAR1--PAXBAR30)/30;
SVAR2 = Var(Of PBXBAR1--PBXBAR30)/30;
SVAR3 = Var(Of PCXBAR1--PCXBAR30)/30;
SVAR4 = Var(Of PDXBAR1--PDXBAR30)/30;
SVAR5 = Var(Of PEXBAR1--PEXBAR30)/30;

SVAR = Mean(Of SVAR1-SVAR5);
MVAR = Var(Of SS1-SS5);
SE = Sqrt(SVAR+((6/5)*MVAR));
CV = (SE/XBAR) * 100;

if (cntrid ne .);
if (gender ne .);

Proc Print;
Title 'Country and Gender by Mean Prose Scores';
Var cntrid gender UNW SWT XBAR SE CV;
run;
```

The final output will have two lines for each country, one for males and the other for females. The variable XBAR provides the mean score and the variable SE provides the standard error for the mean estimate. CV provides the coefficient of variation. The variable SWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SAS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program V:

Country and Gender by Mean Prose Scores						
		UNW	SWT	XBAR	SE	CV
Germany	Men	938	26874222	276.8	2.0	0.7
	Women	1124	26952066	274.9	1.4	0.5
United States	Men	1434	75312132	269.3	2.3	0.8
	Women	1601	84983108	277.6	2.2	0.8
Ireland	Men	1077	1092200	262.9	5.3	2.0
	Women	1346	1082180	268.4	2.3	0.9
Netherlands	Men	1358	5325766	281.5	1.3	0.5
	Women	1479	5134592	283.9	1.4	0.5
Poland	Men	1431	12130543	227.9	1.1	0.5
	Women	1569	12345106	231.0	1.7	0.7
Sweden	Men	1289	2667574	300.6	1.8	0.6
	Women	1355	2692237	301.9	1.7	0.6

VI. Standard error and ratio computation for literacy levels: Multiweight method using SAS with correction for imputation

This program computes proportions at each literacy level and the associated standard errors for any of the three literacy scales. The estimates can be produced for any categorical break variable or a combination of categorical break variables. The following example produces the proportion of individuals at each prose literacy level along with their associated standard errors within each country.

```
Data A;
Set libname.filename (keep= cntrid (or other break variable(s)) plev1--plev5 (or dlev1--dlev5
or qlev1--qlev5) age weight replic01--replic30);
if ((age>=16 and age<=65) or age=9);
if plev1=5 then plev1=4;
if plev2=5 then plev2=4;
if plev3=5 then plev3=4;
if plev4=5 then plev4=4;
if plev5=5 then plev5=4;

Proc Summary Data=A;
Class cntrid;
Var weight replic01-replic30;
Output Out=B N(weight)=UNW
Sum(weight replic01-replic30)=
SWT SW1-SW30;
Proc Summary Data=A;
```

{The bold characters are the only variables and strings that need modification}

```
Class cntrid plev1;
Var weight replic01-replic30;
Output Out=C N(weight)=UNW1
Sum(weight replic01-replic30)=
VSWT VSW1-VSW30;
Proc Summary Data=A;
Class cntrid plev2;
Var weight replic01-replic30;
Output Out=D N(weight)=UNW2
Sum(weight replic01-replic30)=
WSWT WSW1-WSW30;
Proc Summary Data=A;
Class cntrid plev3;
Var weight replic01-replic30;
Output Out=E N(weight)=UNW3
Sum(weight replic01-replic30)=
XSWT XSW1-XSW30;
Proc Summary Data=A;
Class cntrid plev4;
Var weight replic01-replic30;
Output Out=F N(weight)=UNW4
Sum(weight replic01-replic30)=
YSWT YSW1-YSW30;
Proc Summary Data=A;
Class cntrid plev5;
Var weight replic01-replic30;
Output Out=G N(weight)=UNW5
Sum(weight replic01-replic30)=
ZSWT ZSW1-ZSW30;

Proc Sort Data=B;
By cntrid;
Proc Sort Data=C(rename=(plev1=plev));
By cntrid plev;
Proc Sort Data=D(rename=(plev2=plev));
By cntrid plev;
Proc Sort Data=E(rename=(plev3=plev));
By cntrid plev;
Proc Sort Data=F(rename=(plev4=plev));
By cntrid plev;
Proc Sort Data=G(rename=(plev5=plev));
By cntrid plev;

Data H nonH ProblemH;
Merge B(in=b) C(in=c) D(in=d) E(in=e) F(in=f) G(in=g);
By cntrid;
if b and c and d and e and f and g then output H;
else if b and not(c) then output nonH;
else if b and not(d) then output nonH;
else if b and not(e) then output nonH;
else if b and not(f) then output nonH;
else if b and not(g) then output nonH;
else if not(b) and c then output problemH;
else if not(b) and d then output problemH;
else if not(b) and e then output problemH;
else if not(b) and f then output problemH;
else if not(b) and g then output problemH;
run;
```

```
Data I;
Set H;
Array SW SW1-SW30;
Array VSW VSW1-VSW30;
Array WSW WSW1-WSW30;
Array XSW XSW1-XSW30;
Array YSW YSW1-YSW30;
Array ZSW ZSW1-ZSW30;
Array AXBAR AXBAR1-AXBAR30;
Array BXBAR BXBAR1-BXBAR30;
Array CXBAR CXBAR1-CXBAR30;
Array DXBAR DXBAR1-DXBAR30;
Array EXBAR EXBAR1-EXBAR30;
Do over SW;
AXBAR = VSW/SW;
BXBAR = WSW/SW;
CXBAR = XSW/SW;
DXBAR = YSW/SW;
EXBAR = ZSW/SW;
end;

XBAR1 = VSWT/SWT;
XBAR2 = WSWT/SWT;
XBAR3 = XSWT/SWT;
XBAR4 = YSWT/SWT;
XBAR5 = ZSWT/SWT;
XBAR = Mean(Of XBAR1-XBAR5);

Array PAXBAR PAXBAR1-PAXBAR30;
Array PBXBAR PBXBAR1-PBXBAR30;
Array PCXBAR PCXBAR1-PCXBAR30;
Array PDXBAR PDXBAR1-PDXBAR30;
Array PEXBAR PEXBAR1-PEXBAR30;
Do over AXBAR;
PAXBAR = (30*XBAR) - (29*AXBAR);
PBXBAR = (30*XBAR) - (29*BXBAR);
PCXBAR = (30*XBAR) - (29*CXBAR);
PDXBAR = (30*XBAR) - (29*DXBAR);
PEXBAR = (30*XBAR) - (29*EXBAR);
end;

SVAR1 = Var(Of PAXBAR1--PAXBAR30)/30;
SVAR2 = Var(Of PBXBAR1--PBXBAR30)/30;
SVAR3 = Var(Of PCXBAR1--PCXBAR30)/30;
SVAR4 = Var(Of PDXBAR1--PDXBAR30)/30;
SVAR5 = Var(Of PEXBAR1--PEXBAR30)/30;

SVAR = Mean(Of SVAR1-SVAR5);
MVAR = Var(Of XBAR1-XBAR5);
SE = Sqrt(SVAR+((6/5)*MVAR));
SVARpct=SVAR*100;
MVARpct=MVAR*100;
XBARpct=XBAR*100;
SEpct= SE*100;
CVpct= (SEpct/XBARpct)*100;
```

if (**cntrid** ne .);
if (**plev** ne .);

Proc Print Data=I;
Title 'COUNTRY BY PROSE LEVEL';
Var **cntrid plev** UNW SWT XBARpct SEpct CVpct;
run;

The final output will have four lines for each country, one for each prose level. The variable XBARpct provides the proportion of individuals at a given level within a country and the variable SEpct provides the standard error for the proportion. CVpct provides the coefficient of variation. The variable VSWT gives a weighted cell count for the first plausible value and UNW1 provides an unweighted cell count for the first plausible value. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SAS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program VI:

COUNTRY BY PROSE LEVEL						
	Prose Scale	UNW1	VSWT	XBARpct	SEpct	CVpct
Germany	Level 1	299	7724776	13.8	1.3	9.5
	Level 2	715	18424267	35.3	1.4	3.9
	Level 3	768	20439707	37.3	1.7	4.5
	Level 4/5	280	7237538	13.6	1.0	7.1
United States	Level 1	843	33117817	20.8	1.2	5.7
	Level 2	778	41438190	24.5	1.5	6.0
	Level 3	881	51895358	32.8	1.3	3.9
	Level 4/5	533	33843876	21.9	1.3	5.9
Ireland	Level 1	557	491423.3	22.5	1.6	7.3
	Level 2	738	647543.1	30.5	1.6	5.3
	Level 3	812	742544	33.8	1.7	5.0
	Level 4/5	316	292869.7	13.2	1.7	12.6
Netherlands	Level 1	256	1101215	10.4	0.7	6.5
	Level 2	812	3147855	29.4	1.0	3.4
	Level 3	1292	4612593	44.7	1.4	3.1
	Level 4/5	477	1598696	15.5	1.1	7.1
Poland	Level 1	1288	10438093	42.7	0.8	1.9
	Level 2	1036	8455395	34.3	0.9	2.6
	Level 3	587	4834112	19.2	0.8	4.1
	Level 4/5	89	748048.4	3.7	0.6	17.1
Sweden	Level 1	192	401835.1	7.2	0.6	8.0
	Level 2	530	1089885	20.7	0.8	3.8
	Level 3	1059	2129211	39.8	1.0	2.5
	Level 4/5	863	1738880	32.2	1.0	3.2

VII. Standard error computation for quantitative variables excluding literacy scores: Multiweight method using SAS.

This program computes standard errors for quantitative variables other than the plausible values (i.e. other than the literacy scores). The mean for variables such as duration of training (Derived by multiplying variables F8M1*F9M1*F10M1, gives duration of first mentioned course/program) or other continuous variables that may be derived can have their means calculated along with the standard error in the following program. The following example produces mean estimates for years of education and their associated standard errors for each gender within each country.

```
Data A;
Set libname.filename (keep= cntrid (or other break variable) gender (or other break
variable(s)) quantitative variable (e.g. years of education – a7) age weight replic01 -- replic30);
if ((age >=16 and age <=65) or age=9);

Array WT replic01-replic30;
Array AWX AWX1-AWX30;
Do Over WT;
AWX = WT*a7;
end;

WS = a7*weight;

Proc Summary Data=A;
Class cntrid gender;
Var weight replic01-replic30 AWX1-AWX30 WS;
Output Out=B N(weight)=UNW
Sum(weight replic01-replic30 AWX1-AWX30 WS)=
SWT SW1-SW30 ASX1-ASX30 SS;

Data C;
Set B;
Array SW SW1-SW30;
Array VSX ASX1-ASX30;
Array AXBAR AXBAR1-AXBAR30;
Do over SW;
AXBAR = VSX/SW;
end;

XBAR = SS/SWT;

Array PAXBAR PAXBAR1-PAXBAR30;
Do over AXBAR;
PAXBAR = (30*XBAR) - (29*AXBAR);
end;

SVAR = Var(Of PAXBAR1--PAXBAR30)/30;
SE = Sqrt(SVAR);
CV = (SE/XBAR) * 100;

if (cntrid ne .);
if (gender ne .);

Proc Print;
Title 'Country and Gender by Years of Education';
Var cntrid gender UNW SWT XBAR SE CV;
run;
```

{The bold characters are the only variables and strings that need modification}

The final output will have two lines for each country, one for males and the other for females. The variable XBAR provides the mean years of education and the variable SE provides the standard error for the mean estimate. CV provides the coefficient of variation. The variable SWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SAS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program VII:

Country and Gender by Years of Education						
		UNW	SWT	XBAR	SE	CV
Germany	Men	931	26734502	11.6	0.1	1.2
	Women	1110	26559411	11.1	0.1	0.8
United States	Men	1355	70826370	13.4	0.1	0.6
	Women	1530	81363215	13.0	0.1	0.6
Ireland	Men	1075	1090292	10.3	0.2	1.6
	Women	1341	1078606	10.5	0.1	0.9
Netherlands	Men	1358	5325766	13.1	0.1	0.8
	Women	1479	5134592	12.2	0.1	0.8
Poland	Men	1430	12120961	11.0	0.1	0.7
	Women	1558	12259003	10.9	0.1	0.6
Sweden	Men	1286	2661256	11.7	0.1	0.7
	Women	1352	2686417	11.6	0.1	0.8

VIII. Standard error computation for background (categorical) variables: Multiweight method using SAS:

This program computes proportions and the associated standard errors using the thirty replicate weights for any two (or more, if additional break variables are added) categorical background variables. The following example produces the proportions of males and females along with their associated standard errors within each country.

```
Data A;
Set libname.filename (keep= cntrid (or other break variable) gender (or other break
variable(s)) age weight replic01--replic30);
if ((age>=16 and age<=65) or age=9);

Proc Summary Data=A;
Class cntrid;
Var weight replic01-replic30;
Output Out=B N(weight)=UNWT
Sum(weight replic01-replic30)=
SWT SW1-SW30;
```

{The bold characters are the only variables and strings that need modification}

```
Proc Summary Data=A;
Class cntrid gender;
Var weight replic01-replic30;
Output Out=C N(weight)=UNW
Sum(weight replic01-replic30)=
ZSWT ZSW1-ZSW30;

Proc Sort Data=B;
By cntrid;
Proc Sort Data=C;
By cntrid gender;

Data D nonD ProblemD;
Merge B(in=b) C(in=c);
By cntrid;
if b and c then output D;
else if b and not(c) then output nonD;
else if not(b) and c then output problemD;
run;

Data E;
Set D;
Array SW SW1-SW30;
Array ZSW ZSW1-ZSW30;
Array AXBAR AXBAR1-AXBAR30;
Do over SW;
AXBAR = ZSW/SW;
end;

XBAR = ZSWT/SWT;

Array PAXBAR PAXBAR1-PAXBAR30;
Do over AXBAR;
PAXBAR = (30*XBAR) - (29*AXBAR);
end;

SVAR = Var(Of PAXBAR1--PAXBAR30)/30;
SE = Sqrt(SVAR);
SVARpct=SVAR*100;
XBARpct=XBAR*100;
SEpct= SE*100;
CVpct= (SEpct/XBARpct)*100;

if (cntrid ne .);
if (gender ne .);

Proc Print Data=E;
Title 'Country by gender';
Var cntrid gender UNW ZSWT XBARpct SEpct CVpct;
run;
```

The final output will have two lines for each country, one for males and the other for females. The variable XBARpct provides the proportion of individuals within each gender type and the variable SEpct provides the standard error for the proportion. CVpct provides the coefficient of variation. The variable ZSWT gives a weighted cell count and UNW provides an unweighted cell count. An error message indicating that a division by zero has been attempted may result since some of the replicate weights are zero. SAS swiftly deals with the problem of dividing by zero, by setting the result to a system missing value and proceeding with the computations without any effect on the final results.

The following is an example of the output produced from program VIII:

Country by Gender						
		UNW	ZSWT	XBARPCT	SEPCT	CVPCT
Germany	Men	938	26874222	49.9	1.2	2.4
	Women	1124	26952066	50.1	1.2	2.4
United States	Men	1434	75312132	47.0	1.1	2.3
	Women	1601	84983108	53.0	1.1	2.0
Ireland	Men	1077	1092200	50.2	1.0	2.0
	Women	1346	1082180	49.8	1.0	2.0
Netherlands	Men	1358	5325766	50.9	0.1	0.2
	Women	1479	5134592	49.1	0.1	0.2
Poland	Men	1431	12130543	49.6	0.0	0.0
	Women	1569	12345106	50.4	0.0	0.0
Sweden	Men	1289	2667574	49.8	1.2	2.4
	Women	1355	2692237	50.2	1.2	2.4

References

KORN and GRAUBARD (1999). *Analysis of Health Surveys*, Wiley.

MURRAY, T.S., KIRSCH, I.S., and JENKINS, L.B. (Eds.) (1998). *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*, National Center for Education Statistics, US Department of Education, Washington DC.

WOLTER, K.M. (1985). *Introduction to Variance Estimation*, Springer-Verlag, New York.

8.2

Non-Sampling errors

Over a large number of observations, randomly occurring non-sampling errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort was made to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized and coding and edit quality checks to verify the processing logic.

Despite these efforts, non-sampling error is bound to exist in every survey. The following text outlines the most likely sources of this error and its impact on the IALS survey.

Sampling Frame:
Once the population for a survey has been defined, the next step is to establish a means to access this population. The sampling frame provides the means. However, there are a number of issues that may arise with respect to the suitability of a frame. One of the main issues in evaluating a frame is the issue of under coverage, where not all elements that should be in the population are on that frame. The coverage for each country was illustrated in Tables 5.1a and 5.1b (Section 5). As the tables showed, all countries achieved a high level of coverage.

Non-response:
A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response.

Total non-response occurred when the interviewer was either unable to contact the respondent, no member of the household was able to provide the information, or the respondent refused to participate in the survey. The non-response rate for the IALS varied by country (See section 6.5). However, analysis of the characteristics of the IALS non-respondents suggests that they are not concentrated in any specific group. Total non-response was handled by adjusting the weight of households who responded to the survey to compensate for those who did not respond.

Partial non-response to the survey occurred, in most cases, when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information. Generally, the extent of partial non-response was small in the IALS. However, one of the variables, which was particularly difficult to collect, as in all surveys, was income.

The IALS had three income questions:

J2: What is the best estimate of your personal income in (year) from all sources, including those just mentioned?

J3: What is the best estimate of your personal income from only wages, salary or self-employment in (year)?

J5: What is the best estimate of the total income of all household members (including yourself) from all sources in (year)?

Other key variables in the IALS are the education questions A5 and A8:

A5: Before you first immigrated to country of interview, what was the highest level of schooling you had completed?

A8: What is the highest level of schooling you have completed?

Response Error:
A number of other potential sources of non-sampling error that are unique to the IALS deserve comment. Firstly, some of the respondents may have found the test portion of the study intimidating and this may have had a negative affect on their performance. Unlike "usual" surveys, the IALS test items have "right" and "wrong" answers. Also, for many respondents this would have been their first exposure to a "test" environment in a considerable number of years. Further, although interviewers did not enforce a time limit for answering questions, the reality of having someone watching and waiting may have, in fact, imposed an unintentional time pressure. Along with these possible response biases, the use of incentives by some countries and the overall level of motivation of respondents between countries might have had an effect, according to some experts, on the scores obtained by respondents. The IALS study team performed numerous analyses in that regard (*Adult Literacy in OECD countries: Technical Report on the first International Adult Literacy Survey*, chapter 6). The study of the possible effect of the use of incentives in other American studies and the effect of motivation (omit rates, not reached rates and time taken on test) on scores obtained by IALS respondents did not show strong and consistent links. Instead, the evidence from IALS strongly suggests that participants' motivation to do well did not bias the study results.

Scoring:
Another potential source of non-sampling error for the IALS relates to the scoring of the test items, particularly those that were scored on a scale (e.g. items that required respondents to write). Special efforts such as centralizing the scoring and sample verification were made to minimize the extent of scoring errors.

8.2.1

Quality Notes

Combining Sample Files Within A Country

Users may want to combine data from different countries or regions in order to obtain results at a more global geographical level. For example, one could combine data for French and English Canada, or data for German, French and Italian regions of Switzerland, or in the United Kingdom, data for Great Britain and Northern Ireland.

The user must take notice of possible limitations in combining these files by consulting the country specific notes that follow.

General Notes

There have been twenty countries that have participated in the IALS survey to date. Each country was responsible for editing their own data file. A generic international record layout (IRL) was provided to each country with instructions on how to create their own national data file. The national files were reviewed by Statistics Canada to search for any deviations from the IRL. This process attempted to identify flow errors, missing categories, and anything out of the ordinary. Despite these efforts, the international data files remain less than perfect. The following provides brief notes on deviations from the IRL that have remained on the data file, by country.

Note that flow errors that involve less than ten to twelve cases have been omitted from the following quality report. These are thought to have a negligible impact on statistical results.

There also exist outlier values in some of the questions involving ranges of acceptable data. Efforts were made to identify and rectify the outliers. Therefore, the extent to which there are outliers is minimal.

Belgium Flanders

Section A

- The 26 responses coded ‘DK/Refused’ (Code 98) in Q.A8 were not flowed into Q.A11 and Q.A12.

Section B

- There were 570 cases in Q.B14, Q.B15, & Q.B17 that were not imputed to the language mentioned in Q.B13L1. The IRL states that if only one language was reported in Q.B13, then Q.B14 to QB17 should be imputed with the same language code.

Section F

- Q.F3M1 to Q.F3M3 should have the same number of responses as Q.F4 to Q.F11.
- Note that there is a high proportion of responses that answered 'Yes' (Code 1) in Q.F6G (i.e. 'Other').
- Q.F12M3 has a total of 89 responses, but according to the flow from Q.F11M3 there should be 119 responses. There are 30 missing responses.

Section J

- In Q.J1A, there are 644 responses coded 'No' (Code 2), but in Q.J3 there are 804 responses coded 'No income' (Code 0).

Canada

General

- Canada's French and English samples **can** be combined and analyzed together without affecting the representation of the populations.
- For users who wish to merge Canada's National file with the international file, please note that the numbering of the questions in Section G (only), beyond Q.G6 is different between the two files. This is due to an additional question that was on the Canadian version of the survey.

Section B

- Q.B2L2 to Q.B5L2 have 174 cases that appear to be missing according to the flow from Q.B1L2. However, when we consider the rule used to determine whether or not the respondent's mother tongue matched the language of interview, the flow is correct. If there were two mother tongues reported and neither Q.B1L1 nor Q.B1L2 had a mother tongue equal to the language of interview, then the responses were flowed into Q.B2L2 to Q.B5L2. There were only 18 responses that satisfied these criteria.
- For Q.B6, and Q.B9 to Q.B12, please read the National Official Language as English if **interview** was in English and French if **interview** was in French. Note that individuals who were interviewed in one language can have performed the literacy tasks in the other language. See notes for BQLANG and TBLANG on the record layout.

Section E

- For Q.E4 to Q.E7, please read the National Official Language as English if **interview** was in English and French if **interview** was in French. Note that individuals who were interviewed in one language can have performed the literacy tasks in the other language. See notes for BQLANG and TBLANG on the record layout.

Section G:

- For Q.G2, Q.G3, Q.G5, Q.G11 to Q.G13, and Q.G15, please read the National Official Language as English if **interview** was in English and French if **interview** was in French. This can cause confusion when we consider the individuals that were interviewed in one language and performed the literacy tasks in the other language. See notes for BQLANG and TBLANG on the record layout.

Chile

General

- No flow errors affecting population estimates were detected.

Czech Republic

Section D

- Differences in the formulation of Q.D1 and the omission of "30 hours" in Q.D5 may have contributed to some consistency problems between Q.D5 and Q.D13.
- Q.D12 was omitted.
- No flow errors affecting population estimates were detected.

Denmark

General

- No flow errors affecting population estimates were detected.

Finland

Section A

- Due to technical differences in the formulation of Q.A8, some of Finland's Q.A8 (ISCED) values have been replaced by data from the Finnish register file. The responses in Q.A7 have not been replaced or adjusted, therefore some of the responses in Q.A7 do not correlate with the data in Q.A8. Similarily, the responses in Q.A9, Q.A11, and Q.A12 have not been replaced, therefore some of the population counts for these variables do not match the current skip patterns out of Q.A8.

General

- No other flow errors affecting population estimates were detected.

Germany

Section A

- Q.A2 has 1,900 responses that were imputed to 'Germany' (Code 28).
- Q.A3 has 2 responses that indicate the respondents immigrated before they were born.
- In Q.A12, there should be 1,302 responses but there are only 131 responses. There are 1,171 responses missing.

Section B

- There is one illegal value in Q.B1L2. The value '0' is not a valid category.

Section E

- Responses should be constant throughout the whole section at 1,237. But Q.E5, Q.E7, & Q.E9 only have 114, 155 & 212 responses respectively.
- For Q.E4 to Q.E9 there should be the same number of responses as in Q.E1 to Q.E3, but the number of responses are not consistent.

Section F

- Germany's survey did not treat this section in a comparable manner to the international survey. They excluded a major component of their adult education process (i.e. apprenticeship training). Therefore, Germany's adult and education training is understated as a result.
- Q.F4AM1 to Q.F4EM1 are missing 130 responses.
- Q.F4AM2 to Q.F4EM2 are missing 64 responses.
- Q.F4AM3 to Q.F4EM3 are missing 31 responses.
- Q.F12M1 has 26 extra responses.

Section G

- Q.G16D has 16 responses missing. The non-responses from Q.G15D were not flowed into this question.

Section J

- In Q.J3, there are 612 responses coded 'No income' (Code 0), but in Q.J1A only 468 responses were coded 'No' (Code2).

Hungary

General

- No flow errors affecting population estimates were detected.

Section J

- Q.J2, QJ3 and Q.J5 - the income reference period for these questions was reformulated to collect monthly estimates; the requirements were for annual estimates.
- Hungary's income quintiles for personal, wage and household income (J2Q, J3Q and J5Q) were calculated using monthly income data. Each country was required to calculate income quintile distributions using external data sources representing annual income estimates for the entire population. The unavailability of reliable income data sources in Hungary has constrained this country to use data collected from this survey to calculate quintile ranges.

Note: See section 5.3 for additional notes on data quality pertaining to Hungary.

Ireland

Section A

- Q.A8 has a total of 2,361 responses, but according to the flow from Q.A7 there should be a total of 2,423 responses. Six responses coded 'No schooling' (Code 0) in Q.A7 were not imputed to 'No schooling' (Code 10) in Q.A8. There are an additional 56 missing responses.
- Q.A12 has a total of 1,301 responses, but according to the flow from Q.A8 the total should be 1,318 responses. There are 17 missing responses.

Section D

- Q.D3 has a total of 1,178 responses, but there should only be 978 responses.
- Q.D4 has a total of 1,317 responses, but according to the flow from Q.D1 and Q.D2 there should be 1,466 responses. There are 149 missing cases.
- Q.D5, Q.D7/Q.D8 (ISCOF), Q.D9 (ISICF) up to Q.D14 should have the same number of responses as Q.D4. However, they vary from 1,387 to 1,431.
- Q.D15 has a total of 380 responses, but according to the flow of the section there should be 417 responses. There are 27 missing responses.
- Q.D16 has a total of 376 responses, but according to the flow from Q.D15 (as it is) there should only be 226. There are 150 extra cases.
- Q.D17 has a total of 162 responses, but according to the flow from Q.D15 (as it is) there should be 203 responses. There are 41 missing responses.
- Q.D19 has a total of 931 responses, but according to the flow from Q.D3 (as it is) there should be 1,178 responses. There are 247 missing responses.
- Q.D21 has a total of 87 responses, but according to the flow from Q.D19 (as it is) there should be 252 responses. There are 165 missing responses.
- Q.D22 has a total of 71 responses, but there should be the same amount of responses as Q.D21.

Section E

- This section should have the same number of responses throughout but it varies from 1,417 to 1,477 responses.

Section F

- The number of responses for Q.F4M1, Q.F4M2, & Q.F4M3 to Q.F14M1, Q.F14M2, & Q.F14M3 should all be constant and equal to the number of responses in Q.F3M1, Q.F3M2, & Q.F3M3 respectively, but the number of responses vary from question to question.

Italy

General

- No flow errors affecting population estimates were detected.

Note: See section 5.3 for additional notes on data quality pertaining to Italy.

Netherlands

Section A

- Q.A12 should have a total of 1,336 responses, but there are only 882 responses. There are 454 responses missing.

Section F

- In Q.F5, there are a high proportion (i.e. 66.29%) of responses coded 'Other' (Code 7).
- In Q.F7, only three of the ten international categories were used. Hence there is a high proportion (i.e. 76.12%) of responses coded 'Elsewhere' (Code 10).

New Zealand

General

- There is a very high proportion of not stated responses in some questions due to the fact that New Zealand gathered limited information via a small questionnaire administered to 922 persons who refused the background questionnaire.

Section A

- The 922 non-response cases from Q.A1 were not flowed into Q.A5.
- The responses coded 'No schooling' (Code 0) in Q.A7 were not imputed to 'No schooling' (Code 10) in Q.A8, rather they were left uncoded in Q.A8.
- Q.A9 has a total of 1,875 responses, but according to flow from Q.A8 there should be 2,392 responses.
- Q.A11 has a total of 1,364 responses, but according to the flow from Q.A8 there should be 1,753 responses. There are 389 missing responses.
- Q.A12 has a total of 1,423 responses but according to the flow from Q.A8 and Q.A11 there should be 1,843 responses. There are 420 missing responses.

Section B

- Q.B13, Q.B14, & Q.B17 are examples of how New Zealand conducted their survey. According to the flow from Q.B1L1 there should be a total of 4,223 responses in these questions. However, only 3,301 responses were flowed in to Q.B13.
  For some questions New Zealand coded the difference with not stated responses (i.e. 922 responses), but for these questions they did not.

Section C

- Q.C1, Q.C5, Q.C7, Q.C8, & Q.C11 are examples of how New Zealand conducted their survey. There are 922 missing responses; For some questions New Zealand coded the difference with not stated responses (i.e. 922 responses), but for these questions they did not.

Section D

- The 922 responses coded 'Not stated' (Code 9) in Q.D2 were not carried into Q.D3 and Q.D4 through the rest of the section.

Section F

- There was 922 cases not carried through this section.

Section J

- There was 922 cases not carried through Q.J4 and Q.J5.
- Q.J3 has 689 responses coded to 'No income' (Code 0), however, Q.J1A has 706 responses coded as 'No' (Code 2).

Norway

Section A

- Due to technical differences in the formulation of Q.A8, Norway's original data collected for Q.A7 and Q.A8 has been replaced by data from the Norwegian register file. The responses to Q.A9, Q.A11 and Q.A12 have been adjusted to reflect the current skip patterns from question Q.A8. This adjustment has resulted in a large number of records being imputed to “not stated” (code 9) in Q.A9 and the suppression of all data in Q.A11.

General

- No other flow errors affecting population estimates were detected.

Note: See section 5.3 for additional notes on data quality pertaining to Norway.

Poland

Section A

- The responses coded ‘No schooling’ (Code 0) in Q.A7 were coded ‘ISCED 0’ (Code 0) in Q.A8 rather than being imputed to ‘No schooling’ (Code 10).
- Q.A10 has a total of 1,319 responses, but according to the flow from Q.A9 there should only be 1,103 responses. There are 219 extra cases.
- Q.A12 has a total of 2,782 responses, but according to the flow from Q.A8 there should only be 1,889 responses. There are 893 extra responses.

Section D

- Q.D15 has a total of 374 responses, but according to the flow from Q.D14 there should only be 350 responses. There are 24 extra cases.

Section J

- Q.J5 has 179 responses not imputed from Q.J2.

Slovenia

General

- A duplicate record (IDNO 127) was detected on the data file. The background questionnaire data is identical whereas the literacy score values are different.
- No flow errors affecting population estimates were detected.

Section J

- Each country was required to calculate income quintile distributions using external data sources representing annual income estimates for the entire population. The unavailability of reliable income data sources in Slovenia has constrained this country to use data collected from this survey to calculate quintile ranges.

Sweden

Section A

- The responses coded ‘No schooling’ (Code 0) in Q.A7 were not imputed to ‘No schooling’ (Code 10) in Q.A8.
- For Q.A5 and Q.A.8, category ‘ISCED 7’ (Code 7) was omitted from the BQ.

Section B

- In Q.B1L1, there are 29 responses coded ‘Refused’ (Code 98). These responses are not carried throughout the rest of the section.

Section C

- For Q.C5 and Q.C.11, category ‘ISCED 7’ (Code 7) was omitted from the BQ.

Section D

- In Q.D3, there are a high proportion (i.e. 24.7%) of responses coded ‘Don’t know/refused’ (Code 98), but no responses are coded ‘Never worked’ (Code 0).
- In Q.D5, there are 37 responses coded ‘Don’t know/refused’ (Code 8) that were not carried to Q.D6.

Section E

- For Q.E1, Q.E2, Q.E3A & Q.E3B, category ‘Once a week’ (Code 3) was omitted from the BQ. They asked either greater or less than once a week.

Section F

- Q.F5 has 1,011 responses but it should have 1,407 responses.
- In Q.F5, categories ‘An apprenticeship certificate’ (Code 4) and ‘Professional or career upgrading’ (Code 6) were omitted from the BQ.
- Q.F8 has 1,006 responses but should have 1,407 responses.
- In Q.F11, the first mention has 0 responses, but the second mention has 556 responses. There should be 1,407 responses in the first mention.

Section G

- For Q.G11, Q.G12, and Q.G13 the category ’Moderate’ (Code 3) was omitted from the BQ.

Switzerland

General

- In Switzerland, the collection period for the German and French regions and the collection period for the Italian region was four years apart. Considering the fact that these three separate samples represented the population of 16 years old and over at the time of collection, combining these population groups into one single group would consist of representing a hybrid population that never existed. This impacts the production of estimates for total population counts but has a lesser impact on the estimation of proportions for that same population. One also needs to be aware of the fact that for Switzerland, test language was geographically dictated. The latter permits us to think that the overall estimated level of literacy skills in the population could have been higher had respondents been given the choice of language to complete the test regardless of their place of residence.

Switzerland French

Section A

- Q.A4 and Q.A5 have a total of 428 responses, but according to the flow from Q.A3 there should only be 340 responses. There are 88 extra responses.
- For Q.A5 and Q.A8, categories 'ISCED 0' (Code 0) and 'ISCED 6' (Code 6) were omitted from the BQ.

Section B

- Q.B4L2 and Q.B5L2 have 95 extra responses according to the flow from Q.B1L2.

Section C

- In Q.C5 and Q.C11, categories 'ISCED 0' (Code 0) and 'ISCED 6' (Code 6) were omitted from the BQ.

Section D

- Q.D12 has a total of 1,041 responses, but according to the flow from Q.D11 there should be 1,171 responses. There are 130 missing responses.

Section F

- Q.F3 has a 100% non-response rate.

Switzerland German

Section A

- Q.A4 and Q.A5 have a total of 278 responses, but according to the flow from Q.A3 there should be 228 responses. There are 50 extra responses.
- For Q.A5 and Q.A8, categories 'ISCED 0' (Code 0) and 'ISCED 6' (Code 6) were omitted from the BQ.

Section B

- Q.B4L2 and Q.B5L2 have 96 extra responses according to the flow from Q.B1L2.

Section C

- For Q.C5 and Q.C11, categories 'ISCED 0' (Code 0) and 'ISCED 6' (Code 6) were omitted from the BQ.

Section D

- Q.D12 has a total of 1,023 responses, but according to the flow from Q.D11 there should be 1,165 responses. There are 142 missing responses.

Section F

- Q.F3 has a 100% non-response rate.

Switzerland Italian

General

- No flow errors affecting population estimates were detected.

Section J

- Each country was required to calculate income quintile distributions using external data sources representing annual income estimates for the entire population. The unavailability of reliable income data sources in Switzerland (Italian) has constrained this country to use data collected from this survey to calculate quintile ranges.

Note: See section 5.3 for additional notes on data quality pertaining to Switzerland (Italian).

United Kingdom

General

- United Kingdom's Great Britain and Northern Ireland samples **can** be combined and analyzed together without affecting the representation of the populations.

United Kingdom Great Britain

Section B

- Q.B5L1 has 88 responses, but according to the flow from Q.B1L1 there should be 254 responses. There are 166 missing responses.
- Q.B14 to Q.B17 have 2,992 responses, but according to the flow of the section there should be 3,811 responses. There are 819 missing responses.

Section D

- Q.D4 has 2,638 responses, but according to the flow from Q.D2 there should be 2,761 responses. There are 123 missing responses.
- The 123 non-responses in Q.D5 are not flowed in to Q.D6.
- Q.D10 has 2,429 responses, but according to the flow from Q.D9 (ISCOR) there should be 2,761 responses. There are 332 missing responses.
- Q.D11 has 2,815 responses, but according to the flow from Q.D9 (ISCOR) there should only be 2,761 responses. There are 54 extra responses.
- Q.D12 and Q.D13 have 2,740 responses, but according to the flow of the section there should be 2,761 responses. There are 21 missing responses.
- Q.D15 has 585 responses, but according to the flow from Q.D14 there should be 608 responses. There are 23 missing responses.
- Q.D16 has 26 missing responses. Q.D17 has 671 total responses. The flow from Q.D15 indicates that the total responses should be 333. All responses that flowed in to Q.D16 should of been directly flowed to section E.

United Kingdom Northern Ireland

Section B

- Q.B12 has 40 missing responses.

Section D

- The 21 non-responses from Q.D5 were not carried into Q.D6.
- The 20 non-responses from Q.D16 were not carried into Q.D17 and Q.D18.

United States

Demographic Section

- The urban/rural values for the variable COMMSIZE have been reversed.

Section A

- In Q.A5 and Q.A8, category 'ISCED 0' (Code 0) was omitted from the BQ.
- Q.A9 has a total of 2,123 responses, but according to the flow from Q.A8 there should be 2,170 responses. There are 47 missing responses.
- Q.A12 has a total of 732 responses, but according to the flow from Q.A8 there should be 887 responses. There are 155 missing responses.

Section B

- For Q.B2L2 to Q.B5L2 there are 16 missing responses according to the flow from Q.B1L2.
- For Q.B18 only one mention was used. This implies that individuals surveyed are classified as belonging to only one ethnic or cultural group, and does not allow for the possibility of mixed ethnic origin.

Section C

- In Q.C5 and Q.C11, categories 'No schooling' (Code10) and 'ISCED 0' (Code 0) were omitted from the BQ.
- Q.D7/Q.D8 (ISCOF), Q.D9 (ISICF) have a total of 2,845 responses, but according to the flow from Q.C11 there should be 2,901 responses. There are 56 missing responses.

Section D

- Q.D6 has a total of 501 responses, but according to the flow from Q.D5 there should be 570 responses. There are 69 missing responses.
- In Q.D6, category 'Other personal or family responsibilities' (Code 3) was omitted from the BQ.
- Q.D21 and Q.D22 have 74 missing responses according to the flow from Q.D19. The non-responses from Q.D19 were not flowed into these questions.

9.0 Questionnaires

To view any of the following questionnaires, activate the hand tool and click on the corresponding country name. Once the questionnaire is opened, you can link back to the user's guide by clicking inside the blue box on the first or last page of the selected questionnaire.

10.0 Record Layouts with Univariate Frequencies

The following section contains the record layouts for each of the 19 IALS countries. On the right-hand margin of the layout are found the unweighted and weighted counts for each variable on the file.

These record layouts should always be consulted when using the microdata files. They contain notes, which will aid in the understanding of the data. Users are cautioned that in many cases the code numbers for variables on the record layout will not correspond with those on the questionnaire for that variable.

As can be observed from the following record layouts, the logical record length of the microdata file is 1,509; the data file contains 479 variables and 64,175 records. The approximate storage space required for the flat file, SPSS file, and SAS file are 95MB 54MB, and 241MB respectively.

To view any of the following record layouts, activate the hand tool and click on the corresponding country name. Once the record layout is opened, you can link back to the user's guide by clicking inside the blue box on the first or last page of the selected record layout.

Appendix 1

International Standard Industrial Classification (ISIC 1968)

Major (10) and sub-major (34) groups

1. Agriculture, hunting, forestry and fishing
Agriculture and hunting
Forestry and logging
Fishing

2. Mining and quarrying
Coal mining
Crude petroleum and natural gas production
Metal ore mining
Other mining

3. Manufacturing
Manufacture of food, beverages and tobacco
Textile, wearing apparel and leather industries
Manufacture of wood and wood products, including furniture
Manufacture of paper and paper products, printing and publishing
Manufacture of chemicals and chemical, petroleum, coal, rubber and plastic products
Manufacture of non-metallic mineral products, except products of petroleum and coal
Basic metal industries
Manufacture of fabricated metal products, machinery and equipment
Other manufacturing industries

4. Electricity, gas and water
Electricity, gas and steam
Water works and supply

5. Construction

6. Wholesale and retail trade, and restaurants and hotels
Wholesale trade
Retail trade
Restaurants and hotels

7. Transport, storage and communication
Transport and storage
Communication

8. Finance, insurance, real estate and business services
Financial institutions
Insurance
Real estate and business services

9. Community, social and personal services

Public administration and defence
Sanitary and similar services
Social and related community services
Recreational and cultural services
Personal and household services
International and other extra-territorial bodies

0. Activities not adequately defined

International Standard Classification of Occupations (ISCO 1988)

Major (10) and sub-major (28) groups

1. Legislators, senior officials and managers
Legislators and senior officials
Corporate managers
General managers

2. Professionals
Physical, mathematical and engineering science professionals
Life science and health professionals
Teaching professionals
Other professionals

3. Technicians and associate professionals
Physical and engineering science associate professionals
Life science and health associate professionals
Teaching associate professionals
Other associate professionals

4. Clerks
Office clerks
Customer services clerks

5. Service workers and shop and market sales workers
Personal and protective services workers
Models, salespersons and demonstrators

6. Skilled agricultural and fishery workers
Market-oriented skilled agricultural and fishery workers
Subsistence agricultural and fishery workers

7. Craft and related trades workers
Extraction and building trades workers
Metal, machinery and related trades workers
Precision, handicraft, printing and related trades workers
Other craft and related trades workers

8. Plant and machine operators and assemblers
Stationary-plant and related operators
Machine operators and assemblers
Drivers and mobile-plant operators

9. Elementary occupations
Sales and services elementary occupations
Agricultural, fishery and related labourers
Labourers in mining, construction, manufacturing and transport

0. Armed forces

Major Field of Study—Final Classification Structure

01 EDUCATIONAL, RECREATIONAL AND COUNSELLING SERVICES (001-009)

001 Education - General
002 Elementary - Primary Education
003 Secondary Education (Basic)
004 Secondary Education (Specialized)
005 Special Education
006 Non-teaching Educational Fields
007 Physical Education, Health and Recreation
008 Counselling Services and Personal Development
009 Other Education

02 FINE AND APPLIED ARTS (010-016)

010 Fine Arts
011 Music
012 Other Performing Arts
013 Commercial and Promotional Arts
014 Graphic and Audio-visual Arts
015 Creative and Design Arts
016 Other Applied Arts

03 HUMANITIES AND RELATED FIELDS (017-026)

017 Classics, Classical and Dead Languages
018 History
019 Library and Records Science
020 Mass Media Studies
021 English Language and Literature
022 French Language and Literature
023 Other Languages and Literature
024 Philosophy
025 Religious Studies
026 Other Humanities and Related Fields

04 SOCIAL SCIENCES AND RELATED FIELDS (027-039)

027 Anthropology
028 Archeology
029 Area Studies (Non Languages or Literature)
030 Economics
031 Geography
032 Law and Jurisprudence
033 Man/Environment Studies
034 Political Science
035 Psychology
036 Sociology
037 Social Work and Social Services
038 War and Military Studies
039 Other Social Sciences and Related Fields

05 COMMERCE, MANAGEMENT AND BUSINESS ADMINISTRATION (040-045)

040 Business and Commerce
041 Financial Management
042 Industrial Management and Administration
043 Institutional Management and Administration
044 Marketing, Merchandising, Retailing and Sales
045 Secretarial Science - General Fields

06 AGRICULTURAL AND BIOLOGICAL SCIENCES/TECHNOLOGIES (046-056)

046 Agricultural Science
047 Agricultural Technology
048 Animal Science Technologies
049 Biochemistry
050 Biology
051 Biophysics
052 Botany
053 Household Science and Related Fields
054 Veterinary Medicine/Science
055 Zoology
056 Other Agricultural and Biological Sciences/Technologies

07 ENGINEERING AND APPLIED SCIENCES (057-070)

057 Architecture and Architectural Engineering
058 Aeronautical and Aerospace Engineering
059 Biological and Chemical Engineering
060 Civil Engineering
061 Design/Systems Engineering
062 Electrical/Electronic Engineering
063 Industrial Engineering
064 Mechanical Engineering
065 Mining, Metallurgical and Petroleum Engineering
066 Resources and Environmental Engineering
067 Engineering Science
068 Engineering, n.e.c.
069 Forestry
070 Landscape Architecture

08 ENGINEERING AND APPLIED SCIENCE TECHNOLOGIES AND TRADES (071-082)

071 Architectural Technology
072 Chemical Technology
073 Building Technologies
074 Data Processing and Computer Science Technologies
075 Electronic and Electrical Technologies
076 Environmental and Conservation Technologies
077 General and Civil Engineering Technologies
078 Industrial Engineering Technologies
079 Mechanical Engineering Technologies
080 Primary Industries/Resource Processing Technology
081 Transportation Technologies
082 Other Engineering/Applied Science Technologies, n.e.c.

09 HEALTH PROFESSIONS, SCIENCES AND TECHNOLOGIES (083-098)

083 Dentistry
084 Medicine - General
085 Medicine - Basic Medical Science
086 Medical Specializations (Non-surgical)
087 Paraclinical Sciences
088 Surgery and Surgical Specializations
089 Nursing
090 Nursing Assistance
091 Optometry
092 Pharmacy and Pharmaceutical Sciences
093 Public Health
094 Rehabilitation Medicine
095 Medical Laboratory and Diagnostic Technology
096 Medical Treatment Technologies
097 Medical Equipment and Prosthetics
098 Other Health Professions, Sciences and Technologies, n.e.c.

10 MATHEMATICS AND PHYSICAL SCIENCES (099-109)
099 Actuarial Science
100 Applied Mathematics
101 Chemistry
102 Geology and Related Fields
103 Mathematical Statistics
104 Mathematics
105 Metallurgy and Materials Science
106 Meteorology
107 Oceanography and Marine Sciences
108 Physics
109 General Science

11 ALL OTHER N.E.C. (110)
110 All Other - (Not Elsewhere Classified)

12 NO SPECIALIZATION (111)
111 No Specialization

13 NO POSTSECONDARY QUALIFICATION (112)
112 No Postsecondary Qualification

14 UPGRADING (120-130)
120 Upgrading - General
121 Basic Education (Grades 1-8)
122 General Education (G.E.D. - High School equivalency)
123 High School Subjects (Secondary Credit, Grades 9-13)
124 Post Secondary Upgrading
125 Pre-Vocational Upgrading
126 Basic Training for Skill Development (B.T.S.D.)
127 Basic Job Readiness Training (B.J.R.T. - job entry program)
128 Orientation
129 Career Alternatives (Job Hunting)
130 University Transfer

15 PERSONAL DEVELOPMENT (131-138)
131 Personal Development - General
132 Home and Family
133 Consumer/Financial
134 Coping Skills
135 Communications Skills
136 Religion and Morals
137 Public Affairs, Community/Current Events
138 Driver Instruction

16 RECREATIONAL ACTIVITY (139-141)
139 Sports and Outdoor Recreation
140 Physical Fitness
141 Games

The following are special codes that apply only to Great Britain and Northern Ireland.

Major Field of Study (Major Headings)—Final Classification Structure

150 Educational, Recreational and Counselling Services
151 Fine and Applied Arts
152 Humanities and Related Fields
153 Social Sciences and Related Fields
154 Commerce, Management and Business Administration
155 Agricultural and Biological Sciences/Technologies
156 Engineering and Applied Sciences
157 Engineering and Applied Science Technologies and Trades
158 Health Professions, Sciences and Technologies
159 Mathematics and Physical Sciences
160 All Other N.E.C. (Not Elsewhere Classified)
161 No Specialization
162 No Postsecondary Qualification
163 Upgrading
164 Personal Development
165 Recreational Activity